

TARTU ÜLIKOOL
LOODUS- JA TÄPPISTEADUSTE VALDKOND
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT
BIOTEHNOLOOGIA ÕPPETOOL

Madli Tamm

**Kasvajaseoseliste mutatsioonide määramine kopsuvähiga
uuritavate rakuvabast DNA-st**

Magistritöö (30 EAP)

Juhendaja	MD, PhD Neeme Tõnisson
Kaasjuhendajad	PhD Paula Ann Kivistik
	MSc Mart Kals

TARTU 2016

Kasvajaseoseliste mutatsioonide määramine kopsuvähiga uuritavate rakuvabast DNA-st

Magistritöö käigus töötati välja kopsuvähispetsiifiliste mutatsioonide tuvastamiseks alleel-spetsiifilise fragmentanalüüsi ja amplikonipõhise teise põlvkonna sekveneerimise töövoog. Neid meetodeid kasutades analüüsiti 70 kopsuvähiga patsiendi plasmast eraldatud rakuvaba DNA-d ja 33 FFPE koelõigust eraldatud genoomset DNA-d. Nimetatud meetoditega testides on võimalik tuvastada mutantset alleeli, mille esinemissagedus on alla 1%. FFPE DNA analüüsimise tundlikkus on kõrgem võrreldes rakuvaba DNA-ga, kuid seevastu rakuvaba DNA spetsiifilisus on suurem.

Märksõnad: tsirkuleeriv rakuvaba DNA, kasvajaseoselised mutatsioonid, kopsuvähk, teise põlvkonna sekveneerimine.

B520 Üldpatoloogia, patoloogiline anatoomia.

Detecting cancer related mutations in cell-free DNA of lung cancer patients

An allele-specific fragment length analysis (FLA) and an amplicon-based multiplex next-generation sequencing (NGS) workflow for lung cancer-related genes was set up. Plasma samples from 70 lung cancer patients with 33 matching FFPE materials were analysed using these methods. Results show that optimized workflows for FLA and NGS approaches can detect mutant allele content with the frequency of less than 1%. The sensitivity of analysing FFPE DNA material was better than that of cfDNA, but cfDNA had better specificity over FFPE DNA material.

Key words: cell-free DNA, cancer related mutations, lung cancer, next-generation sequencing.

B520 General pathology, pathological anatomy.

SISUKORD

KASUTATUD LÜHENDID	5
SISSEJUHATUS	6
KIRJANDUSE ÜLEVAADE.....	7
Onkogenees	7
Kopsuvähki põhjustavad keskkonnategurid	9
Kopsuvähi histoloogilised vormid.....	10
Kopsuvähi juhtmutatsioonid ja ravitundlikkus.....	10
Kopsuvähi diagnostika	14
Vedel biopsia	15
Molekulaardiagnostilised meetodid onkoloogias	17
EKSPERIMENTAALOSA	20
Töö eesmärgid	20
Materjal ja metoodika.....	21
Analüüsimetoodikate väljatöötamine ja optimeerimine	21
Rakuvaba ja koematerjali DNA eraldamine	21
DNA kvaliteedi kontroll	22
Fragmentanalüüs.....	22
TruSight® Tumor 26 paneel.....	25
Amplikonipõhine teise põlvkonna sekveneerimine.....	26
Uuritavad	33
Uuritavate rakuvaba ja koelõigu DNA	35
Ravimseoseliste mutatsioonide testimine	35
Mutatsioonileidude võrdlemine	36
Tulemused	37
Analüüsimetoodikate väljatöötamine ja optimeerimine	37
Rakuvaba ja FFPE DNA eraldamine ning kvaliteedi kontroll	37
Fragmentanalüüs.....	38
TruSight® Tumor 26 paneel.....	40
Amplikonipõhine teise põlvkonna sekveneerimise	42
Uuritavad	46
Uuritavate cfDNA ja FFPE DNA eraldamine	46
Ravimseoseliste mutatsioonide leiud.....	46
Arutelu	54
KOKKUVÕTE	56
Summary.....	58
TÄNUAVALDUSED.....	59

KASUTATUD KIRJANDUS	60
Kasutatud veebiaadressid	66
LISAD	67
Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks	71

KASUTATUD LÜHENDID

bp base pair, aluspaar

cfDNA *cell-free* DNA, rakuvaba DNA

EGFR *epidermal growth factor receptor*, epidermaalne kasvufaktori retseptor

FFPE *formalin-fixed, paraffin-embedded*, formaliinis fikseeritud, parafiini sisestatud

NGS *next-generation sequencing*, teise põlvkonna sekveneerimine

NSCLC *non-small cell lung cancer*, mitteväikerakuline kopsukartsinoom

SISSEJUHATUS

Vananeva rahvastiku tõttu on vähkkasvajatesse haigestumus arenenud maades suurenemas. Nii vähi tekkel kui ka progressioonil ja ravimresistentsusel on oluline roll koesisestel geneetilistel muutustel. Tänu sellele arusaamale on hakatud lisaks kasvaja koe histoloogilisele iseloomustamisele kasutatama ka molekulaardiagnostilisi meetodeid määramaks ravitundlikke mutatsioone. Viimati nimetatud analüüsi teeb keeruliseks kasvajarakkude heterogeensus ja koe biopsia võtmisel saadav vähene analüüsitav materjal. Biopsiamaterjalist on võimalik tuvastada ainult konkreetses asukohas konkreetsel ajahetkel esinevaid mutatsioone. Nimetatud kitsaskohti saaks lahendada „vedela biopsia“ kasutuselevõttuga. Vedelaks biopsiaks nimetatakse üldjuhul verest eraldatud biomarkereid, näiteks DNA molekule. Kuna kehavedelikku satub DNA molekule kõikidest kudedest, on kasvajast pärit mutantse DNA osakaal madal. Sellest tulenevalt on tsirkuleerivate DNA molekulide genotüpiseerimiseks vajalik kasutada piisavalt tundlikke meetodeid.

Antud magistritöö üheks eesmärgiks oligi välja töötada metoodika rakuvaba DNA molekulaarseks analüüsiks ning seejärel testida väljatöötatud metoodika kasutatavust uuringus osalevate kopsuvähiga patsientide koematerjali ja rakuvaba DNA peal leidmaks ravimtundlikke mutatsioone.

Eeldatavalt muutub hetkel veel valdavalt siirdemeditiini valdkonda kuuluva tsirkuleeriva rakuvaba DNA analüüs kliinilises onkoloogias peatselt igapäevaselt kasutatavaks. Käesolev magistritöö püüab selleks pakkuda efektiivseid metoodilisi lahendusi.

KIRJANDUSE ÜLEVAADE

Onkogenees

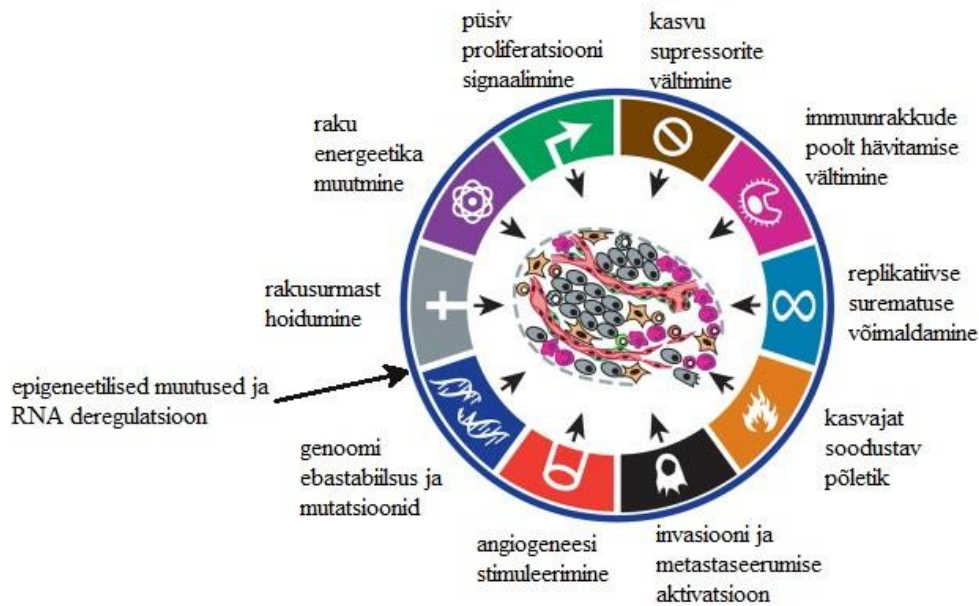
Vähkkasvaja on defineeritud kui geneetiline haigus, mida põhjustab selliste mutatsioonide kuhjumine, mis soodustavad rakkude kontrollimatut paljunemist, invasiivseid omadusi ja metastaseerumist (Cooper, 2005, Devita *et al.*, 2015a, Fearon, 1997). Kuigi enamus tuumorites esinevatest mutatsioonidest on somaatilised, põhjustades sporaadilisi kasvajaid, on siiski umbes 1% kõigist kopsukasvajatest pärilikud (Cooper, 2005).

Tuumorigeneesi põhjustavad muutused kolme tüüpi geenides: onkogeenides, kasvaja supressor geenides ja stabiilsus geenides. Esimesel juhul mutatsioonid nendes geenides põhjustavad pidevat aktivatsiooni või aktiveerumist tingimustel, kui metsiktüüpi (*wild-type*) geen oleks vaigistatud, millele vastupidiselt muteerunud tuumori supressor geenilt kodeeritud produkti aktiivsus on pärssitud (Vogelstein & Kinzler, 2004). Onkogeenides esinevad mutatsioonid peamiselt spetsiifilistes kuumades punktides (*hot spot*), mis on valdavalt missens- ehk aminohappe asendusmutatsioonid (Devita *et al.*, 2015a). Onkogeeni ühes alleelis esinev aktiveeriv somaatiline mutatsioon on enamasti piisav, et põhjustada raku selektiivset kasvueelset (Vogelstein & Kinzler, 2004). Tuumori supressor geenides esineb seevastu inaktiveerivaid mutatsioone kogu geeni ulatuses. Pahaloomulistes kasvujates esinevad peamiselt nukleotiidi vahetused, indelid (insertsioonid ja deletsioonid), kromosomaalsed ümberkorraldused ning koopiaarvu muutused (Devita *et al.*, 2015a). Kolmas klass kasvaja teket soodustavaid geene, mida kutsutakse stabiilsus või *caretaker* geenideks, aitavad muteerudes tuumorigeneesile kaasa hoopis teistmoodi. Siia klassi kuuluvad valepaardumisreparatsiooni (*mismatch repair*, MMR), nukleotiidi väljalõike (*nucleotide excision repair*, NER) ja lämmastikaluse väljalõike (*base excision repair*, BER) geenid, mille produktid vastutavad normaalsel DNA replikatsioonil või mutageenidega kokkupuutel tekkinud vigade parandamise eest. Samuti kuuluvad siia geenid, mille produktid kontrollivad suuremaid kromosoomi alasid hõlmavaid protsesse, mis vastutavad näiteks korrektse mitootilise rekombinatsiooni ja kromosomaalse segregatsiooni eest. Stabiilsus geenid hoiavad seega geneetiliste muutuste arvu madalal ning juhul, kui need geenid on inaktiveeritud, siis toimub muteerumine teistes geenides kõrgemal määral (Friedberg, 2003).

Lisaks geneetilistele mutatsioonidele on vähitekkes oluline roll ka epigeneetilistel muutustel (Cooper, 2005, Kok-Sin *et al.*, 2015). Siia kuuluvad muutunud metüülatsioonitase, histoonide kovalentsed modifikatsioonid ja mikro-RNA-de (miRNA) poolt vahendatud geeniregulatsiooni muutused. Kasvaja teket võivad soodustada üldine hüpometüülatsioon, mis

põhjustab genoomset ebastabiilsust, ja tuumor supressorgeenide hüpermetülatsioon, mis põhjustab nende geenide vaigistamist. Anormaalne metülatsioon võib mõjutada ka mõningate miRNA-de ekspressiooni (Brambilla & Gazdar, 2009). DNA metülatsiooniga on seotud ka nukleosoomi histooni sabade modifikatsioonid, mistõttu on muutunud kromatiini pakkimine ja geeniekspressioon (Fraga *et al.*, 2005). Tuumor supressor- ja onkogeenidena võivad käituda ka miRNA-d (Calin *et al.*, 2004). Normaalsetes rakkudes reguleerivad miRNA-d märklaud mRNA-de ekspressiooni (Jackson & Standart, 2007), kuid vähkkasvajate puhul on üldjuhul miRNA-de biogenees düsreguleeritud (Hata & Lieberman, 2015). Võib väita, et miRNA-d on seotud kasvaja patogeneesiga, kuna need osalevad rakkude vohamisel ja apoptoosil (Brennecke *et al.*, 2003, Xie *et al.*, 2016) ning miRNA-de geenid asuvad sageli habrastes saitides ja genoomsetes regioonides, mida seostatakse vähkkasvajatega (Calin *et al.*, 2004).

Erinevalt paljudest haigustest, nagu näiteks tsüstiline fibroos ja lihasdüstroofia, mida põhjustab mutatsioon ühes geenis, on vähkkasvajatele iseloomulik see, et ainult ühe geeni defekt ei põhjusta kasvajad. Imetaja rakkudes on mitmeid kontrollpunkte, mis kaitsevad tekkinud potentsiaalsete kahjulike mutatsioonide eest ning seetõttu ongi vaja mitme geeni defekti, et areneks invasiivne tuumor (Vogelstein & Kinzler, 2004). Kasvaja kujunemine on astmeline protsess, mille käigus omandab tuumor mitmeid joonisel 1. toodud bioloogilisi võimeid (Hanahan & Weinberg, 2011). Tuumori tekke evolutsioonilise protsessi käigus toimub mutatsioonide kuhjumine erinevates rakuliinides, mistõttu on rakkudevaheline heterogeensus paratamatu. See omadus teebki kasvaja koe molekulaargeneetilise analüüsi keerukaks (Cooper, 2005).



Joonis 1. Hanahani ja Weinbergi poolt välja pakutud bioloogilised tunnused, mille on kas kõik või suur enamus kasvaja arengu käigus omandanud (Hanahan & Weinberg, 2011). Lisaks nende poolt pakutule võib 11. omaduseks lugeda ka epigeneetilised muutused ja RNA deregulatsiooni (Imielinski *et al.*, 2012).

Kopsuvähki põhjustavad keskkonnategurid

Kopsuvähi tekkel on oluline nii geneetiline taust kui ka keskkonnategurid. Geneetilisi häireid soodustavateks keskkonnateguriteks on:

- suitsetamine ja passiivne suitsetamine;
- asbest;
- radioaktiivne tolm ja radooni aurud.

Ligikaudu 85% kõigist kopsuvähi juhtudest on seotud suitsetamisega (Padrik & Everaus, 2013). Kopsuvähi tekke tõenäosuse ja suitsetamise doosi vahel esineb tugev seos (Wang *et al.*, 2015). Päevas paki sigaretti suitsetaval inimesel on kahekümnekordne risk kopsuvähi haigestuda võrreldes mittesuitsetajaga (Alberg & Samet, 2003, Padrik & Everaus, 2013). Kuna kopsuvähk on suuresti põhjustatud suitsetamisest, siis esineb haige süüdistamist ning madalat prioriteeti kopsuvähi teadusuuringutes (Devita *et al.*, 2015a). Samas üle poole Ameerika Ühendriikides kopsuvähi diagnoosi saanutest on kas mittesuitsetajad või aastate eest suitsetamise maha jätnud (Garces *et al.*, 2004, Tong *et al.*, 1996).

Lisaks nimetatud mõjuritele loetakse kopsuvähi riskifaktoriteks veel õhusaastet (Pope *et al.*, 2002), osooni (Palli *et al.*, 2004), toitumist (Miller *et al.*, 2004), papilloomiviiruse infektsiooni (Cheng *et al.*, 2001) ja kroonilist põletikku (Ballaz & Mulshine, 2003).

Kopsuvähi histoloogilised vormid

Kopsu pahaloomulised kasvaja­jad võivad areneda kopsu parenhüümist või bronhiaalsest koest. Tulenevalt erinevustest kasvaja­te bioloogias ja ravist jagatakse kopsuvähk kahte suuremasse gruppi: mitteväikerakuline kopsukartsinoom (*non-small cell lung cancer*, NSCLC) ja väikerakuline kopsukartsinoom (*small cell lung cancer*, SCLC) (Padrik & Everaus, 2013). Umbes 20-25% kopsuvähi juhtudest on väikerakulised ning 75-80% mitteväikerakulised kasvaja­jad (Cooper, 2005). Mitteväikerakulise kopsuvähi sagedasemad histoloogilised vormid on adenokartsinoom ning lame- ja suurerakuline kopsukartsinoom¹. Kopsuvähile on iseloomulik suur mutatsioonide esinemise hulk võrreldes teiste enamlevinud pahaloomuliste kasvaja­tega (Kandath *et al.*, 2013). Seda võib seletada suitsetajate vahetu kokkupuutega kartsinogeensete ainetega (Govindan *et al.*, 2012). Sellelegi vaatamata on umbes 10-20% kopsuvähihaigetest mittesuitsetajad (Garces *et al.*, 2004, Subramanian *et al.*, 2007). Siia gruppi kuuluvad kõige sagedamini naised adenokartsinoomse histoloogiaga (Devarakonda *et al.*, 2015). Põhiliselt metastaseerub kopsuvähk kaela ja rindkere lümfisõlmedesse, kuid samuti ka rinnakelmesse, maksa, neerupealistesse ja luudesse (Cooper, 2005).

Eesti Tervise Arengu Instituudi tervisestatistika ja terviseuuringute andmebaasi² andmetel oli 2013. aastal Eestis 834 kopsu, bronhi ja hingetoru (C33-C34) pahaloomulise kasvaja esmasjuhtu. Nimetatud paikmete viie aasta suhteline elulemus nii Eestis³ kui kogu Euroopas oli aastatel 1999-2014 kõigest 13%, mis on üks madalamate seast võrreldes teiste kudede vähkkasvaja­tega (Francisci *et al.*, 2015). Kõrge suremus on põhjustatud varajase diagnoosimise keerukusest ja kaugelearenenud kasvaja kiirest progressioonist (Devita *et al.*, 2015b).

Kopsuvähi juhtmutatsioonid ja ravitundlikkus

Soliidtuumoritele omaselt on ka kopsuvähk molekulaarselt heterogeenne haigus (Chen *et al.*, 2013). Mitmed geenid osalevad nii SCLC kui ka NSCLC tuumorigeneesis, kuid sellegi poolest seostatakse iga kasvaja tüübiga unikaalseid geneetilisi muutuseid (Devita *et al.*, 2015b). Nende molekulaarsete anomaalsuste olemuste ja sageduste kindlakstegemine on vajalik määramaks nende kliinilist tähtsust (näiteks seotust suitsetamise, histoloogilise tüübi, kasvaja staadiumi, elulemuse ja ravitundlikkusega), kasutatavust kopsuvähi ennetamise ja varajase diagnoosimise vallas ning lisaks ka ravimiarenduses (Devita *et al.*, 2015b). On

¹ <http://www.cancer.org/>

² <http://pxweb.tai.ee/esf/pxweb2008/dialog/statfile2.asp>

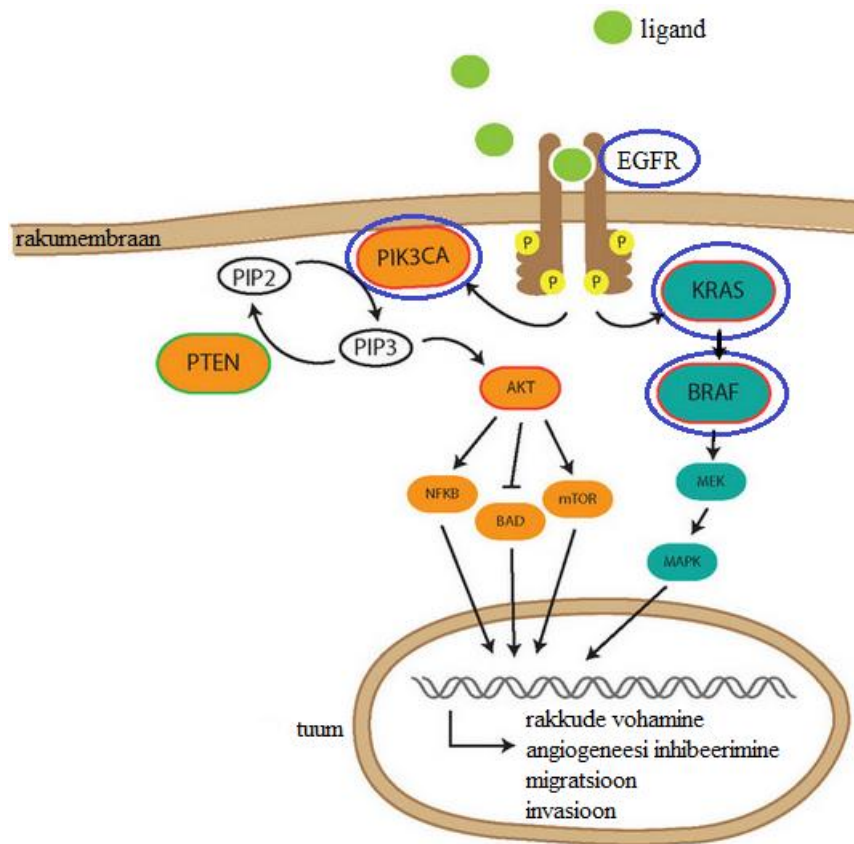
³ <http://www.tai.ee/et/tegevused/registrid/vahiregister>

oletatud, et kasvaja molekulaarse iseloomustamise põhjal oleks võimalik igale patsiendile soovitada personaalselt talle sobiv raviplaan, mis võiks maksimeerida ravi efektiivsust ja vähendada soovimatuid kõrvaltoimeid (Macconail et al., 2011).

Vähkkasvajate ravitundlikkus on määratud konkreetsetes kasvajas esinevate mutatsioonide poolt. Mitmete kasvajate puhul on määratud juhtmutatsioonid geenides, mis kodeerivad signaalvalke, mis on olulised rakkude vohamisel ja ellujäämisel. Umbes 50% kopsu adenokartsinoomidest esineb somaatilisi onkogeenseid mutatsioone EGFR (*epidermal growth factor receptor*), ERBB2 (*erb-b2 receptor tyrosine kinase 2*), KRAS (*Kirsten rat sarcoma viral oncogene homolog*), PIK3CA (*phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha*), BRAF (*B-Raf proto-oncogene, serine/threonine kinase*), MAP2K1 (*mitogen-activated protein kinase kinase 1*) ja ALK (*anaplastic lymphoma receptor tyrosine kinase*) geenides (Su et al., 2011). Toetudes patogeneetilistele signaaliradadele, on välja töötatud mitmeid sihtmärkravimeid, mis häirivad rakkude vohamist (kasvufaktorite inhibiitorid), rakutsükli edasiminekut ja angiogeneesi. Kasvufaktoreid mõjutavad ravimid toimivad peamiselt EGFR ja ERBB2 geeniproduktidele ning KRAS onkogeeni aktivatsioonile. Rakutsükli ja angiogeneesi pidurdavad ravimid toimivad tuumor supressorgeenidele (Huber & Stratakis, 2004).

Epidermaalse kasvufaktori retseptorit (EGFR) kodeeriv geen, mis asub 7. kromosoomi lühikeses õlas, on üks enam esinev muteerunud proto-onkogeen kopsu adenokartsinoomides. See geen kodeerib transmembraanset ERBB perekonna retseptor-türosiinkinaasi, mis aktiveerub, kui selle ekstratsellulaarsele ligandi siduvale regioonile seondub üks EGF perekonna molekulidest. Aktiveerudes need retseptorid homo- või heterodimeriseeruvad teiste ERBB perekonnaliikmetega ning vallandavad RAS/RAF/MEK ja PI3K/AKT/mTOR signaalraja. Epidermaalse kasvufaktori (*epidermal growth factor*, EGF) retseptori perekonda kuuluvat valku kodeerib ka ERBB2 geen, mis asub 17. kromosoomi pikas õlas (Devita et al., 2015b, Hynes & Lane, 2005, Riese & Stern, 1998, Yarden & Sliwkowski, 2001). Valgul endal puudub ligandi siduv domeen, mistõttu see ei seo kasvufaktoreid, vaid seondub tugevalt teiste ligandi siduvate EGF retseptori perekonnaliikmetega, moodustades heterodimeere (Graus-Porta et al., 1997). Moodustunud kompleks seob ligande stabiilselt ja võimendab kinaas-vahendatud allavoolu signaalradasid, kuhu kuuluvad mitogeen-aktiveeritud proteiinkinaas (MAPK) ja fosfatidüülinoositol-3 kinaas (PI3K). Mutatsioonid ja üleekspressioon EGFR ja ERBB2 geenides põhjustavad türosiinkinaasi pidevat aktivatsiooni ning järgnevate radade fosforülatiooni, mille tulemusel soodustub rakkude

kontrollimatu proliferatsioon, invasioon ja metastaaside teke (joonis 2)⁴ (Devita *et al.*, 2015b, Hynes & Lane, 2005, Yarden & Sliwkowski, 2001).



Joonis 2. Signaalirada, mis algab ligandi seondumisega EGFR-le, mistõttu retseptor homodimeriseerub ja fosforüleerub (joonisel tähistatud P-tähega). Aktiveeritud EGF retseptor aktiveerib KRAS onkogeeni, mis omakorda BRAF onkogeeni ja kaks proteiin kinaasi (MEK ja MAPK). Selle tulemusel hakatakse ekspresseerima kasvu soodustavaid gene. EGFR aktiveerib ka PIK3CA onkogeeni, mis fosforüleerib PIP2-e PIP3-ks. Toimub mitmete allavoolu asuvate efektorite aktiveerimine, mis viib omakorda valkude sünteesini, raku kasvamise, vohamise, migratsiooni, angiogeneesi ja invasioonini. Sinise ringiga on tähistatud antud töös käsitletavat geenid (Berg & Soreide, 2012).

KRAS geen asub 12. kromosoomi lühikeses õlas ja kuulub RAS geenide perekonda. See geen kodeerib valku, mis on liige väikeste GTP-aaside superperekonnas. Aktiveerivaid mutatsioone võivad põhjustada ühenukleotiidsed asendused, mida on seostatud mitmete pahaloomuliste kasvajatega, nagu näiteks kopsu adenokartsinoom ja kolorektaalne vähk⁵.

BRAF geen asub 7. kromosoomi pikas õlas ja kodeerib valku, mis kuulub seriin/treoniin proteiin kinaasi raf/mil perekonda. See valk reguleerib rakkude jagunemist, diferentseerumist ja sekretsiooni mõjutavat MAK kinaasi/ERK signaalirada. BRAF geeni mutatsioone seostatakse ka südame defektidest tulenevate haigustega, vaimse arengu peetusega ning

⁴ <http://www.ncbi.nlm.nih.gov/gene/2064>

⁵ <http://www.ncbi.nlm.nih.gov/gene/3845>

mitmete pahaloomuliste kasvajatena, nagu näiteks mitte-Hodgkini lümfoom, melanoom, mitteväikerakuline kopsuvähk ja kopsu adenokartsinoom⁶.

PIK3CA geen asub 3. kromosoomi pikas õlas. Fosfatidüülinositol 3 kinaas (*phosphatidylinositol-3-kinase*, PI3K) koosneb väiksemast regulatoorsest ja suuremast katalüütilisest alaüksusest, mis on kodeeritud erinevate geenide poolt⁷. Regulatoorne subühik säilitab katalüütilise alaüksuse madalat aktiivsust mittepaljunevates (*quiescent*) rakkudes (Cantley, 2002). PIK3CA geen kodeerib kinaasi p110α katalüütilist alaüksust, mis kasutab ATP-d (*adenosine triphosphate*, adenosintrifosfaat), et fosforüleerida lipiide, mille koostises on inositol (fosfatidüülinositoolid, *phosphatidylinositols*, PtdIns). PI3K rada reguleerib rakkude vohamist, kasvu, apoptoosi ja tsütoskeleti ümberkorraldusi. PIK3CA on proto-onkogeen, milles esinevad aktiveerivad mutatsioonid peamiselt p110α helikaalses või kinaasi domeenis ja neid seostatakse erinevate vähkkasvajate tuumorigeneesiga (Samuels *et al.*, 2004).

Umbes 15-20% kauglearenenud mitteväikerakulise kopsuvähiga patsientidel esineb sihtmärkravis kasutatavaid juhtmutatsioone (Bordi *et al.*, 2015). Kaks kõige enam levinud mutatsiooni EGFR geenis on 21. eksoni punktmutatsioon (rs121434568) ja lühikesed raamisisesed (*in-frame*) deletsioonid eksonis 19 (Shigematsu *et al.*, 2005). Neid võib nimetada ravitundlikeks mutatsioonideks, kuna alluvad ravile türosiinkinaasi inhibiitoritega (*tyrosine kinase inhibitor*, TKI) (Devita *et al.*, 2015b), mis pidurdavad EGFR signalisatsiooni ja tulemusena tekib kasvajakraku apoptoos (Devarakonda *et al.*, 2015). Kindlaks tehtud EGFR mutatsiooniga patsientidele kirjutatakse TKI ravimina välja kas gefitinib, erlotinib või afatinib (Bordi *et al.*, 2015).

KRAS ja EGFR mutatsioonide koosesinemise ja prognoosi ning ravitundlikkuse kohta on andmed vastukäivad. On näidatud, et EGFR/KRAS mutatsioonide puhul on progressioonivaba ja üldine elulemus madalam kui patsientidel, kellel on ainult EGFR mutatsioonid (Wang *et al.*, 2010). Samas Kim *et al.* tööühm näitas, et KRAS mutatsioonid ei mõjutanud prognoosi (Kim *et al.*, 2013).

EGFR-TKI ravi kasutamisel on vältimatu omandatud resistentsuse kujunemine, mis tekib keskmiselt 10-12 kuu jooksul (Sequist *et al.*, 2011). Kõige sagedamini tekitab omandatud resistentsust sekundaarne EGFR mutatsioon 20. eksonis (p.T790M, c.2369C>T, rs121434569), mille esinemissagedus jääb vahemikku 49...63% (Sequist *et al.*, 2011, Yu *et*

⁶ <http://www.ncbi.nlm.nih.gov/gene/673>

⁷ <http://www.ncbi.nlm.nih.gov/gene/5290>

al., 2013). Harvemini põhjustavad resistentsuse tekkimist ERBB2 geeni amplifikatsioonid (umbes 12% juhtudest), MET geeni amplifikatsioonid (umbes 8% juhtudest), PIK3CA mutatsioonid (umbes 5%) või BRAF geeni mutatsioonid (1%) (Bordi *et al.*, 2015). Resistentsusmehhanismi identifitseerimine on aluseks raviplaani muutmisele. Äsja on heakskiidu kliiniliseks kasutuseks saanud kolmanda põlvkonna EGFR-TKI ravim osimertinib, mida kasutatakse p.T790M mutatsiooniga patsientide raviks (Greig, 2016).

Kopsuvähi diagnostika

Kopsuvähi diagnostika algab anamneesist, läbivaatusest ja radioloogilistest uuringutest. Kuid kasvaja histoloogiliseks, immuunohistokeemiliseks ja molekulaarseteks uuringuteks on hetkel vaja koematerjali, mille saamiseks kasutatakse:

- bronhoskoopiat, mis võimaldab hinnata kasvaja ulatust ning võtta biopsia ja tsütoloogiline uuring;
- nõelbiopsiat kompuutertomograafia kontrolli all, mis võimaldab võtta biopsiat perifeersetest kasvajatest;
- mediastinoskoopiat, mida kasutatakse kasvaja levikuulatuse hindamiseks mediastinaalsetes lümfisõlmedes;
- operatsiooniaegset biopsia võtmist (Padrik & Everaus, 2013).

Kasvaja suuruse ja levikuulatuse kirjeldamiseks kasutatakse standardiseeritud TNM-süsteemi (*tumor, node, metastasis*, kasvaja, lümfisõlm, metastaas). TNM-süsteem klassifitseerib pahaloomulised kasvajakud algkolde levikuulatuse ja suuruse (T), regionaalsete lümfisõlmede haaratuse (N) ja kaugmetastaaside puudumise või olemasolu (M) alusel (Devita *et al.*, 2015a, Padrik & Everaus, 2013). Samuti määratakse histopatoloogiline diferentseerumisaste ja jääkkasvaja olemasolu ning suurus. Sarnase prognoosiga TNM-grupid koondatakse staadiumitesse 0 (kartsinoom *in situ*), I, II, III ja IV (kaugmetastaasid), mis omakorda jaotatakse alajaotusteks (Ia, Ib jne) (Padrik & Everaus, 2013).

2015. aastal WHO (World Health Organization) poolt välja antud uus kopsu kasvajate klassifikatsioon sisaldab juhiseid, mis soovivad kõikide kasvajatüüpide puhul kasutada diagnoosimisel immuunohistokeemiat (Travis *et al.*, 2015). Tänu nimetatud tehnikale on võimalik NSCLC jagada adeno- ja lamerakulisteks kartsinoomideks. Esimesele neist on iseloomulik TTF-1 (*thyroid transcription factor-1*) esinemine ning teisele p63 (*transformation-related protein 63*), CK (*cytokeratin*) 5 ja 6 (Rekhtman *et al.*, 2011). Lisaks

sellele on vajalik mitteväikerakulise kopsuvähiga patsientide geneetiline testimine, kuna paljud ravimid on efektiivsed ainult kindlate mutatsioonide olemasolul (Travis *et al.*, 2015).

Sõeluuringute ning spetsiifiliste sümptomite puudumine haiguse algfaasis tingib kopsuvähi hilist diagnoosi (Paci *et al.*, 2009). Ameerika Vähiliidu poolt läbiviidud uuringust selgus, et peaaegu pooltel juhtudel on diagnoosi hetkeks vähkkasvaja kaugelearenenud (Siegel *et al.*, 2016). Lisaks sellele on praeguste diagnoosimeetodite puuduseks asjaolu, et biopsia võtmine võimaldab määrata ainult antud ajas ja kohas esinevaid muutuseid. Samuti on invasiivsete omaduste tõttu biopsia teostamine mõnikord vastunäidustatud. Seepärast on viimasel ajal suurt tähelepanu pööratud mitteinvasiivsete uurimismeetodite arendamisele (Ma *et al.*, 2015).

Vedel biopsia

Vedelaks biopsiaks nimetatakse kehavedelikest (näiteks veri, sülg, uriin, kõhuõõne- ja pleuravedelik) isoleeritud biomarkereid (Rolfo *et al.*, 2014). Nendeks võivad olla eksosoomid, tsirkuleerivad kasvaja rakud (*circulating tumor cells*, CTC) või tsirkuleerivad nukleiinhapped (*circulating nucleic acids*, CNA). Vähiuuringute siirdemeditsiinis on hetkel vedel biopsia kiiresti arenev ning laialdaselt uuritud valdkond, kuna see võib olla vähkkasvajaga patsientidel kasutoovaks nii diagnoosil, ravi valikul kui ka haiguse jälgimisel (Rolfo *et al.*, 2014). Vedelat biopsiat saab kasutada:

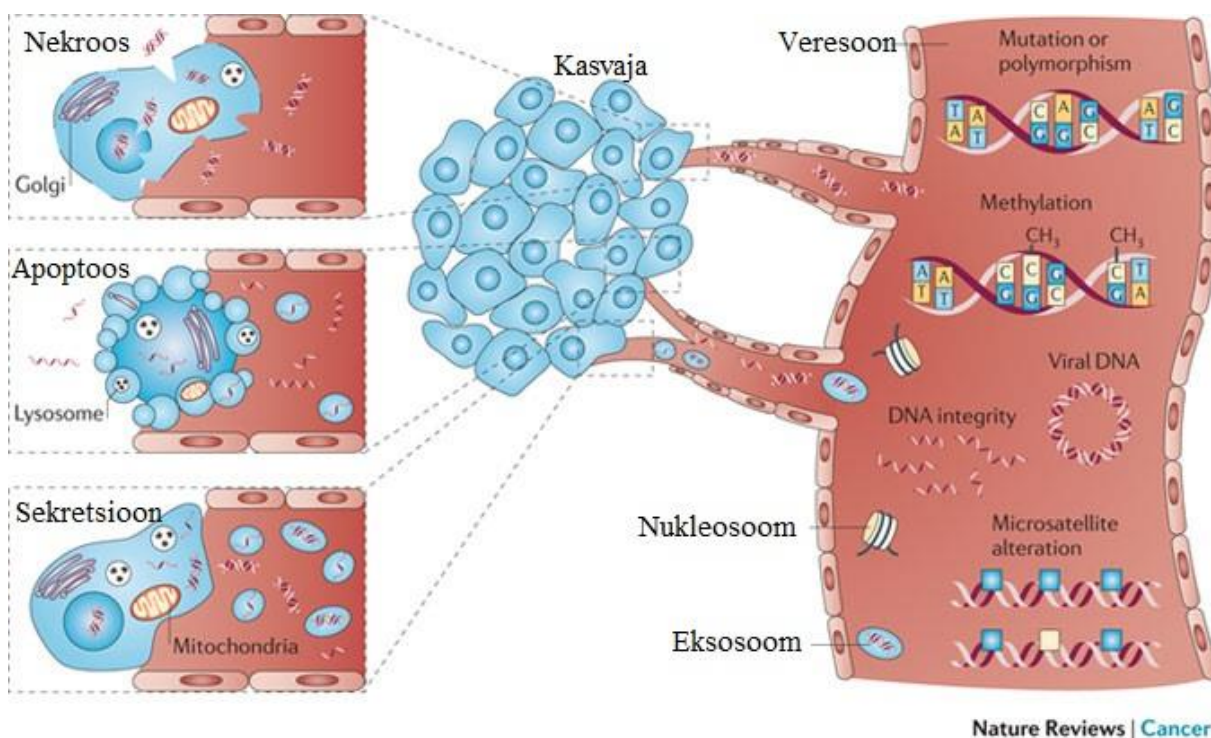
- varajaseks diagnoosimiseks;
- terapeutiliste sihtmärkide ja resistentsuse määramiseks;
- reaalajas ravi tulemuslikkuse jälgimiseks;
- metastaatiliste retsidiivide esinemise või progresseerumise hindamiseks (Rolfo *et al.*, 2014, Schwarzenbach *et al.*, 2011).

Eksosoomide vabanemine eksotsütoosi teel toimub nii normaalse füsioloogilise seisundi (näiteks rakkude omavaheline suhtlus) kui ka kasvaja esinemise korral. Tuumori invasiivsuse ja eksosoomide vabastamise vahel esineb positiivne korrelatsioon. Arvatakse, et vähirakud, millel on suurenenud metastaatiline potentsiaal, sekreteerivad eksosoomi, mis sisaldavad distinktiivsete tunnustega valke. Rakkude migratsiooni stimuleerivaid eksosoomi saab kasutada terapeutilise sihtmärgi või metastaasi indikaatorina (Harris *et al.*, 2015).

Kasvaja massist eemaldudes satuvad rakud lümfi- või vereringesse, kus need võivad jääda ringlema üksikute rakkudena, kobarasse koguneda või kasvada sisse uude koesse, et moodustada metastaase (Gold *et al.*, 2015). Mitteväikerakulise vähkkasvajaga indiviididel on CTC üldarv veres positiivses korrelatsioonis kasvaja staadiumiga (Chen *et al.*, 2015, Krebs *et*

al., 2011), kuid sellegipoolest on patsientidel milliliitris veres keskmiselt kõigest 1...2 CTC-d (Hofman *et al.*, 2011, Zhang *et al.*, 2016).

Tsirkuleerivad nukleiinhapped on ekstratsellulaarsed nukleiinhapped, mida leidub nii tervete kui ka vähkkasvajaga indiviidide kehavedelikes (Swaminathan & Butt, 2006). CNA-deks on nii DNA (*cell-free* DNA, cfDNA ja *circulating tumor* DNA, ctDNA), mRNA kui ka miRNA molekulid (Calin *et al.*, 2004). Rakuvaba kaheahelaliste DNA fragmentide sattumine vereringesse toimub kasvaja rakkude nekroosi või apoptoosi tulemusel (Jahr *et al.*, 2001) ning elusatest rakkudest aktiivse sekretsiooni kaudu (Stroun *et al.*, 2001) (joonis 3).



Joonis 3. Veres ringlevate nukleiinhapete päritolu (Schwarzenbach *et al.*, 2011).

Mitteväikerakulise kopsuvähiga patsientidel on keskmiselt plasma cfDNA kontsentratsioon kõrgem kui tervetel isikutel (Sozzi *et al.*, 2003), olles mõlemas grupis siiski väga varieeruv ning osaliselt kattuv (Szpechcinski *et al.*, 2016). Lisaks sellele võib cfDNA taseme muutus kehavedelikes kajastada ka füsioloogilisi ja patoloogilisi protsesse, mis ei ole kasvajaspetsiifilised (Atamaniuk *et al.*, 2004, Jiang & Pisetsky, 2005, Lo *et al.*, 2000). Vere- ja lümfiringes olevat cfDNA kogust mõjutavad ka füsioloogiline filtreerimine ja degradatsioon. Maks ja neerud eemaldavad cfDNA molekule vereringest, kus nende poolestusaeg varieerub 15 minutist kuni mitmete tundideni (Schwarzenbach *et al.*, 2011). Seetõttu ei saa ainult cfDNA kontsentratsiooni mõõtes diagnoosida kasvaja olemasolu. Vaatamata sellele saab cfDNA molekule kasutada vähkkasvaja biomarkerina, kui tuvastada tuumoritele omaseid geneetilisi ja epigeneetilisi muutuseid (Schwarzenbach *et al.*, 2011).

Tsirkuleeriv rakuvaba DNA võib sisaldada nii kodeerivat kui ka mittekodeerivat osa genoomsest DNA-st, mida saab testida mikrosatelliitse ebastabiilsuse, heterosügootsuse kao (*loss of heterozygosity*, LOH), mutatsioonide, polümorfismi, metülatsiooni ja fragmentide suuruse suhtes (Ma *et al.*, 2015).

Võrreldes koest proovi võtmisega, on vedelal biopsial mitmeid eeliseid: vähene invasiivsus, terviklik pilt kasvaja erinevates kolletes esinevatest mutatsioonidest ja reaajas muutuste jälgimine. Rakuvabast DNA-st mutatsioonide tuvastamise suurimaks probleemiks on piisava spetsiifilisuse ja tundlikkuse tagamine, kuna PCR-il põhinevad tehnoloogiad eeldavad, et muteerunud fragmente on kehavedelikes piisavas hulgas võrreldes normaalsest koest pärinevate DNA järjestustega. Ühtse biomarkerite molekulaargeneetilise testimise standardi puudumise tõttu on keeruline võrrelda omavahel erinevate laborite poolt läbiviidud analüüse (Schwarzenbach *et al.*, 2011). Kuid vaatamata sellele püüdsid Luo *et al.* metaanalüüsi käigus määrata rakuvaba DNA diagnostilist kasutatavust EGFR mutatsioonide tuvastamisel. 20 kriteeriumitele vastava uuringu analüüsimisel arvutati cfDNA-l põhineva mutatsiooni detekteerimise tundlikkuseks 67%, spetsiifilisuseks 94%, šansside suhteks (*diagnostic odds ratio*, DOR) 30 ja kõveraaluseks pindalaks (*area under curve*, AUC) 93%. Viimase kahe kõrge väärtuse põhjal võib öelda, et cfDNA on diagnostikas kasutatav eeskätt siis, kui koematerjali ei ole saada, kuigi tundlikkus jääb koematerjali analüüsile alla (Luo *et al.*, 2014).

Molekulaardiagnostilised meetodid onkoloogias

Onkoloogias vajatakse molekulaardiagnostilisi uuringuid kasvajate varajasel avastamisel, ravi määramisel ja sekundaarsel preventatsioonil, kuna kasvaja tekkega kaasnevad alati geneetilised ja epigeneetilised muutused (Huber & Stratakis, 2004). Diagnostilised laborid testivad hetkel kuumade punktide (*hot spot*) mutatsioone standardsete molekulaarsete tehnoloogiatega. Näiteks kasutatakse kvantitatiivset qPCR-i, mida mõnikord kombineeritakse kõrglahutusega sulamispunkti analüüsiga (*high resolution melting analysis*, HRM), ja Sangeri sekveneerimist. Olgugi et need meetodid on usaldusväärsed, on nende puuduseks suhteliselt suure DNA algkoguse vajadus, suutmatus tuvastada mutatsioone muteerunud järjestuste madala osakaalu puhul ja pikk töötlemise aeg, mis tuleneb vajadusest iga huvipakkuvat piirkonda eraldi testida (Froyen *et al.*, 2016).

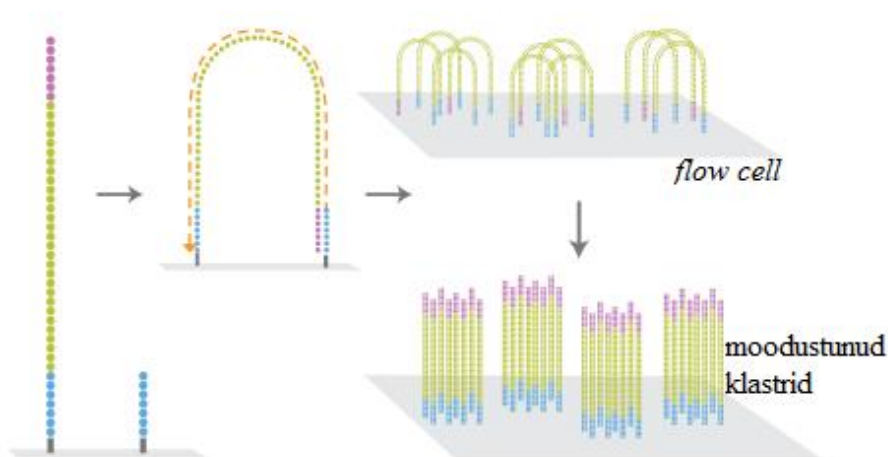
Neid meetodeid sobiks asendada teise põlvkonna sekveneerimine (*next generation sequencing*, NGS) (Froyen *et al.*, 2016), mida saab kasutada nii kogu genoomi, eksoomi kui ka spetsiifilistelt piirkodandelt saadud amplikonide analüüsiks. NGS-i kasutades saab tuvastada muutusi nii DNA, mRNA, miRNA kui ka DNA metülatsiooni tasemel (Devita *et*

al., 2015a). Multipleks amplikonipõhisel genotüpiseerimisel on olenemata väikesest algmaterjali kogusest võimalik saavutada järjestuste suurem katvus, tänu millele saab leida madala alleelse fraktsiooniga ühenukleotiidseid muutusi ning insertioone-deletsioone (indeleid). Amplikonide sekveneerimine on kasulik juhtudel, kui soovitakse tuvastada somaatilisi mutatsioone keerulistest proovidest, nagu näiteks kasvajakud, kus lisaks tuumori DNA-le on ka iduteeliini DNA-d⁸.

Antud magistritöös kasutati huvipakkuvate lühemate regioonide analüüsimiseks Illumina® poolt välja töötatud sünteesil põhinevat sekveneerimist (*sequencing-by-synthesis*, SBS) (Bentley *et al.*, 2008). NGS töövoog neli üldist sammu on:

- raamatukogu valmistamine;
- klastrite loomine;
- sekveneerimine;
- tulemuste analüüs.

Valmistatud adapterjärjestustega üheaahelaline raamatukogu laetakse klaasist kiibile (*flow cell*), kus on füüsiliselt eraldatud rajad. Iga rada on tihedalt kaetud pinnale kinnitatud kahte tüüpi oligotega (p5 ja p7), mis on adapterjärjestustega komplementaarsed. Klastrite loomiseks hakatakse DNA molekule klonaalselt paljundama „silla“ amplifikatsiooni (*bridge-amplification*) meetodil (joonis 4)⁸ (Bentley *et al.*, 2008).



Joonis 4. Sihtmärk DNA klastrite tootmise põhimõte (Illumina, Inc.)

Illumina® SBS tehnoloogia kasutab miljonite moodustunud klastrite samaaegselt sekveneerimiseks ühe aluse kaupa pöörterminaatoritel põhinevat DNA sünteesi meetodit. Iga sekveneerimistsükli käigus liidetakse DNA ahelasse üks fluorestsentsiga märgitud dNTP (A,

⁸http://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf

C, T või G). Klastrite laseriga ergastamise tulemusel tuvastatakse pildistamise tehnikat kasutades fluorestsentsvärv ja selle tugevus ning pärast ensümaatilist lõikamist saab toimuda järgmise nukleotiidi ahela koosseisu lülitumine. Tulemuste analüüsimiseks järjestatakse saadud sekvents referentsjärjestusele⁸ (Bentley *et al.*, 2008).

Lisaks Illumina[®] poolt pakutavale tehnoloogiale on kasutatud kasvajate rakuvaba DNA spetsiifiliste regioonide sekveneerimiseks ka Ion Torrent PGM[™] platvormi (Couraud *et al.*, 2014, Frenel *et al.*, 2015).

EKSPERIMENTAALOSA

Töö eesmärgid

Käesolev magistritöö on osa suuremast „Tsirkuleeriv DNA kopsuvähi biomarkerina“ projektist, mille käigus analüüsitakse mitteväikerakulise kopsuvähiga patsientide vereplasmas ringlevat rakuvaba DNA-d ravimtundlikkust ja prognoosi mõjutavate somaatiliste mutatsioonide ning pikemas plaanis DNA metülatsoonitaseme muutuste suhtes.

Antud töö praktilise osa eesmärgiks oli vereplasmas tsirkuleeriva rakuvaba DNA ja formaliinis fikseeritud, parafiini sisestatud (*formalin-fixed, paraffin-embedded*, FFPE) koematerjalist eraldatud kasvaja DNA testimine sagedamini muteerunud geenides ravimtundlikkust mõjutavate mutatsioonide suhtes, alaeesmärkidega:

1. töötada välja alleel-spetsiifilisel amplifikatsioonil ja fragmentanalüüsil põhinev PCR metoodika EGFR geeni sagedaste mutatsioonide analüüsiks;
2. töötada välja teise põlvkonna sekveneerimisel põhinev meetod sagedasemate ravimtundlikkust ja prognoosi mõjutavate somaatiliste mutatsioonide tuvastamiseks;
3. võrrelda fragmentanalüüsi ja teise põlvkonna sekveneerimise meetodil tuvastatud mutatsioonileide:
 - 3.1. FFPE koelõikudest kliinilistes laborites teostatud EGFR mutatsioonianalüüside tulemustega;
 - 3.2. FFPE koelõikudest ja vereplasmas tsirkuleerivast rakuvabast DNA-st ning nende esindatust ehk alleelseid fraktsioone.

Materjal ja metoodika

Analüüsimetoodikate väljatöötamine ja optimeerimine

Metoodikate väljatöötamiseks ja optimeerimiseks kasutati analüüsitava materjalina tervetelt vabatahtlikelt kogutud vereplasmast eraldatud rakuvaba DNA-d ja Solis BioDyne poolt kloneeritud EGFR geeni 19. eksoni deletsiooni (rs121913421) ja 21. eksoni punktmutatsiooni (rs121434568). Vältimaks alleelide juhuslikku väljalangemist, segati tervete vabatahtlike cfDNA-d omavahel kokku. Meetodite tundlikkuse hindamiseks lisati kontrollide rakuvaba DNA segule erinevates protsentides eelnevalt amplifitseeritud mutantseid DNA fragmente. Samuti kontrolliti metoodikate üldist kasutatavust FFPE koematerjalil.

Rakuvaba ja koematerjali DNA eraldamine

Rakuvaba DNA (*cell-free* DNA, cfDNA) saamiseks koguti tervetelt vabatahtlikelt veeniverd BD Vacutainer® K2E (EDTA) vaakumkatsutitesse (Becton, Dickinson and Company, USA). Vereplasma eraldamiseks tsentrifuugiti verd 1600 rcf (*relative centrifugal force* ehk *gravitational force*, g) juures 15 minutit toatemperatuuril Eppendorf™ Centrifuge 5810 R masinas (Fisher Scientific Company L.L.C., USA), misjärel teostati plasmale teine tsentrifuugimine 2500 rcf juures 10 minutit toatemperatuuril Eppendorf™ Centrifuge 5430 R masinas (Fisher Scientific Company L.L.C., USA) (Wong *et al.*, 2013). Vereplasma 2 ml alikvoodid säilitati -80 °C juures kuni edasise DNA eraldamiseni.

Rakuvaba DNA eraldamiseks plasmast kasutati kahte selleks mõeldud kitti: NucleoSpin® Plasma XS (Macherey-Nagel GmbH & Co. KG, Saksamaa) ja QIAamp® Circulating Nucleic Acid (CNA) (QIAGEN, Saksamaa). Enne cfDNA eraldamist tsentrifuugiti vereplasmat kolmas kord 13000 rcf juures 2 minutit toatemperatuuril Eppendorf™ Centrifuge 5430 R masinas. NucleoSpin® Plasma XS kiti puhul kasutati DNA eraldamiseks modifitseeritud protokoll (lisa 1). QIAamp® CNA kiti puhul kasutati 6 ml plasmast DNA eraldamiseks samuti modifitseeritud protokoll (lisa 2.), kuid 2 ml algmaterjali korral kasutati tootja poolt koostatud juhendit, selle erinevusega, et elueeriti soojendatud elueerimispuhvriga kaks korda.

Metoodikate optimeerimiseks kasutatud FFPE koelõigud saadi SA Tartu Ülikooli Kliinikumi Patoloogiateenistusest. Koematerjalist DNA (FFPE DNA) eraldamiseks kasutati GeneRead™ DNA FFPE (QIAGEN, Saksamaa) kitti vastavalt tootjapoolsetele juhistele, kuid soojendatud puhvriga kaks korda DNA-d elueerides.

DNA kvaliteedi kontroll

Eraldatud DNA kontsentratsiooni mõõdeti Qubit[®] 2.0 fluoromeetriga (Thermo Fisher Scientific Inc., USA) kasutades cfDNA puhul Qubit[®] dsDNA HS (High Sensitivity) Assay (Thermo Fisher Scientific Inc., USA) ja FFPE DNA puhul Qubit[®] dsDNA BR (Broad-Range) Assay (Thermo Fisher Scientific Inc., USA) kitti, vastavalt tootja poolt antud juhtnööridele.

Nii cfDNA kui ka FFPE DNA fragmentide pikkust mõõdeti TapeStation 2200 instrumendil (Agilent Technologies, USA), kasutades lähtuvalt DNA kontsentratsioonist D1000 või High Sensitivity D1000 ScreenTape-i[®] (Agilent Technologies, USA), vastavalt tootja instruksioonidele. Saadud tulemused visualiseeriti TapeStation (Agilent Technologies, USA) tarkvaraga.

Fragmentanalüüs

Fragmentanalüüsi meetodil sooviti tuvastada kahte enamlevinud EGFR geeni mutatsiooni. EGFR geeni 19. eksoni c.2235_2249del15 deletsiooni tuvastamiseks vajalikud praimerjärjestused (*forward* 5'-ACTCTGGATCCCAGAAGGTGAG-3' ja *reverse* 5'-FAM-CCACACAGCAAAGCAGAACTC-3') saadi SA Tartu Ülikooli Kliinikumi Ühendlabori geneetikakeskusest. Nimetatud praimerid paljundavad 121 bp pikkust DNA järjestust. Kuna c.2235_2249del15 deletsioon on 15 nukleotiidi pikk, siis paljundatud mutantse produkti pikkus peab olema 106 bp. EGFR geeni 21. eksoni c.2573T>G punktmutatsiooni tuvastamiseks disainiti käsitsi erinevad alleel-spetsiifilised praimerite komplektid, millest eelkatsete järgselt valiti põhjalikumaks testimiseks kaks *forward* praimerit (*common* 5'-AACACCGCAGCATGTCATGA-3' ja 858R 5'-GTCAAGATCACAGATTTTGGGAG-3') ning üks *reverse* (5'-TAMRA-CTTCTGCATGGTATTCTTTCTCTTC-3') praimer. 858R *forward* oligonukleotiidi puhul jäi otsitav punktmutatsioon praimer 3' otsa alla, tänu millele paljundati ainult mutantset järjestust ning teine, *common forward* praimer disainiti mutatsioonita DNA järjestusele amplifitseeritavuse kontrolliks. Nimetatud praimerid paljundavad 67 bp pikkust mutantset ja 80 bp pikkust referentsjärjestusega produkti.

Määramaks meetodi tundlikkust, tehti tervete kontrollisikute cfDNA segusse erineva hinnangulise koopiarvu suhtega mutantse DNA fragmendi lahjendusrida (*spike-in* 0,2%, 0,5%, 1%, 2%, 4%, 8%). Tervetelt vabatahtlikelt kogutud plasmast eraldatud rakuva DNA-de segu kasutati negatiivse kontrollina, millele lisati kunstliku positiivse ehk mutantse kontrollmaterjali tekitamiseks EGFR geeni 19. eksoni deletsiooni või 21. eksoni punktmutatsiooni sisaldavat amplifitseeritud DNA fragmente.

EGFR geeni 19. eksoni deletsiooni amplifitseerimiseks kasutati kõigepealt reaktsioonisegu, mille kogumaht oli 25 µl ning mis sisaldas lõppkontsentratsioonis 1U HOT FIREPol® DNA Polymerase (Solis BioDyne), 1x HOT FIREPol® Buffer B (Solis BioDyne), 0,2 mM dATP, dTTP, dGTP, dCTP (Thermo Fisher Scientific Inc.), 2,5 mM MgCl₂ (Thermo Fisher Scientific Inc.), 10 pmol praimereid ja *spike-in* lahjendusrida. Amplifikatsioon viidi läbi Applied Biosystems™ 2720 termotsükleris (Thermo Fisher Scientific Inc.) järgmistel tingimustel:

polümeraasi aktivatsioon	95 °C 15 min	
denaturatsioon	95 °C 30 sek	} 40 tsükli
praimerite seondumine	59 °C 30 sek	
elongatsioon	72 °C 60 sek	
lõplik elongatsioon	72 °C 7 min	

programmi lõpp, temperatuuri hoidmine 4 °C.

EGFR geeni 21. eksoni punktmutatsiooni amplifitseerimiseks kasutati kõigepealt reaktsioonisegu, mille kogumaht oli 20 µl ning mis sisaldas lõppkontsentratsioonis 1x HOT FIREPol® GC Master Mix-i (Solis BioDyne), 1 mM MgCl₂ (Solis BioDyne), 6 pmol *reverse* praimerit, 0,6 pmol *common forward* praimerit, 5,4 pmol 858R *forward* praimerit ja *spike-in* lahjendusrida. Amplifikatsioon viidi läbi Applied Biosystems™ 2720 termotsükleris järgmistel tingimustel:

polümeraasi aktivatsioon	95 °C 15 min	
denaturatsioon	95 °C 45 sek	} 16 tsükli
praimerite seondumine	68 °C 40 sek (-0,5 °C tsükli kohta)	
elongatsioon	72 °C 30 sek	
denaturatsioon	95 °C 45 sek	} 30 tsükli
praimerite seondumine	60 °C 30 sek	
elongatsioon	72 °C 30 sek	
lõplik elongatsioon	72 °C 7 min	

programmi lõpp, temperatuuri hoidmine 4 °C.

Fragmentanalüüsi reaktsioon viidi läbi Eesti Biokeskuse tuumiklabori poolt 3130xl Genetic Analyzer (Applied Biosystems, USA) kapillaarelektroforeesi analüsaatoril, mille väljundiks on .fsa laiendiga failid. Saadud andmekogumite analüüsimiseks kasutati GeneMapper® v4.1 tarkvara (Applied Biosystems, USA). Tulemusena saadud elektroferogrammide signaalide ehk piikide põhjal otsustati, kui kõrge on meetodi saagikus, spetsiifilisus ja tundlikkus. Tundlikkuse parandamiseks ja signaalide tugevamaks muutmiseks modifitseeriti mitmeid kordi PCR-i reaktsioonisegu koostist ning amplifikatsiooni tingimusi: denaturatsiooni, praimerite seondumise ja elongatsiooni aega ning temperatuuri ja tsüklite arvu.

Lõplikus protokollis kasutati EGFR geeni 19. eksoni puhul reaktsioonisegu kogumahuga 20 µl, mis sisaldas lõppkontsentratsioonis 1x HOT FIREPol® GC Master Mix-i, 1 mM MgCl₂, 10 pmol praimereid ja *spike-in* lahjendusrida. Amplifikatsioon viidi läbi Applied Biosystems™ 2720 termotsükleris järgmistel tingimustel:

polümeraasi aktivatsioon	95 °C 15 min		
denaturatsioon	95 °C 45 sek	}	16 tsükli
praimerite seondumine	68 °C 40 sek (-0,5 °C tsükli kohta)		
elongatsioon	72 °C 30 sek		
denaturatsioon	95 °C 45 sek	}	36 tsükli
praimerite seondumine	60 °C 30 sek		
elongatsioon	72 °C 30 sek		
lõplik elongatsioon	72 °C 7 min		

programmi lõpp, temperatuuri hoidmine 4 °C.

EGFR geeni 21. eksoni punktmutatsiooni paljundamiseks kasutati lõplikus protokollis reaktsioonisegu kogumahuga 20 µl, mis sisaldas lõppkontsentratsioonis 1x HOT FIREPol® GC Master Mix-i, 1 mM MgCl₂, 10 pmol *reverse* praimerit, 1 pmol *common forward* praimerit, 9 pmol 858R *forward* praimerit ja *spike-in* lahjendusrida. Amplifikatsioon viidi läbi Applied Biosystems™ 2720 termotsükleris järgmistel tingimustel:

polümeraasi aktivatsioon	95 °C 12 min		
denaturatsioon	95 °C 45 sek	}	12 tsüklit
praimerite seondumine	68 °C 40 sek (-0,5 °C tsükli kohta)		
elongatsioon	72 °C 30 sek		
denaturatsioon	95 °C 40 sek	}	20 tsüklit
praimerite seondumine	62 °C 30 sek		
elongatsioon	72 °C 30 sek		
denaturatsioon	95 °C 30 sek	}	26 tsüklit
praimerite seondumine	60 °C 30 sek		
elongatsioon	72 °C 30 sek		
lõplik elongatsioon	72 °C 7 min		

programmi lõpp, temperatuuri hoidmine 4 °C.

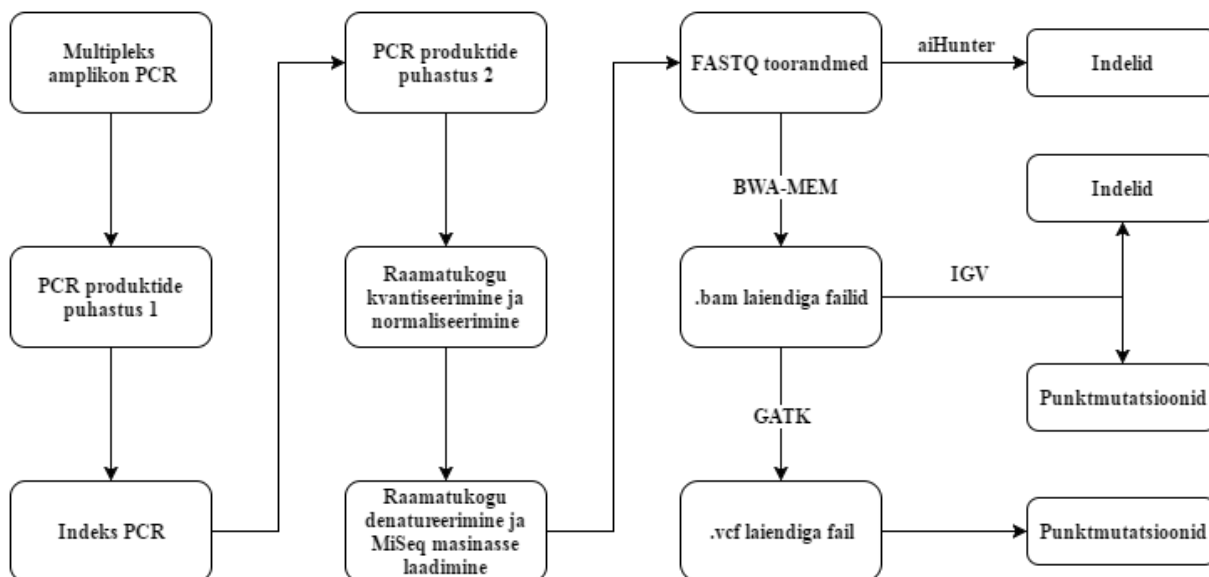
Fragmentanalüüsi reaktsioon ja tulemuste analüüs viidi läbi eelpool kirjeldatud viisil.

TruSight[®] Tumor 26 paneel

Teise põlvkonna sekveneerimise (*next-generation sequencing*, NGS) läbiviimiseks prooviti esmalt Illumina[®] poolt pakutavat TruSight[®] Tumor 26 paneeli, mis on eeskätt mõeldud FFPE koelõikudest eraldatud DNA-st 26 geeni 174 ampliconi sekveneerimiseks ja somaatiliste mutatsioonide tuvastamiseks. Enne raamatukogu koostamist nimetatud paneelile võrreldi nii rakuvaba DNA kui ka FFPE koelõikudest eraldatud DNA amplifitseerumise võimekust *Quality Control Template*-ga (QCT). Illumina[®] protokoll kohaselt tuleb selleks viia läbi proovide 100-kordsete lahjenduste kvantitatiivne reaalaaja PCR (qPCR) 7900HT Fast Real-Time PCR System (Thermo Fisher Scientific Inc.) masinal. Raamatukogu koostamiseks kasutati TruSight[™] Tumor proovide ettevalmistamise protokoll (Illumina, Inc., USA), vastavalt tootja poolt koostatud juhistele, selle erinevusega, etprodukte paljundati 27 tsükli asemel 32 tsükli jooksul. Õige pikkusega raamatukogu olemasolu kontrolliti TapeStation 2200 instrumendil.

Amplikonipõhine teise põlvkonna sekveneerimine

Võttes aluseks Tartu Ülikooli Eesti Geenivaramu tuumiklabori amplikonipõhise teise põlvkonna sekveneerimise protokoll, loodi käesolevas magistritöös multipleks PCR-i põhjal Illumina® platvormil sekveneerimiseks sobiva raamatukogu koostamise ja tulemuste analüüsi töövoog (joonis 5).



Joonis 5. Raamatukogu koostamise töövoog amplikonipõhiseks teise põlvkonna sekveneerimiseks ja tulemuste analüüsiks.

Uuritavate geenide valik ja praimerite disain

Esmalt otsiti kirjandusest ja COSMIC⁹ (Catalogue of somatic mutations in cancer) andmebaasist (Forbes *et al.*, 2015) mitteväikerakulise kopsuvähi puhul 5 kõige sagedasemalt muteerunud ravimseoselist geeni. Valitud geenide eksonitele disainiti praimerid tingimustel, et nn sulamistemperatuur (*melting temperature*, T_m) jääks 60 °C lähistele ning produktide pikkus vahemikku 80...100 bp (tabel 1). Selleks kasutati nii Primer-BLAST programmi (Ye *et al.*, 2012) kui ka käsitsi disainimist. Välistamaks sagedaste pärilike geenivariantide esinemist praimerite seondumiskohtades (*minor allele frequency*, $MAF > 1\%$) teostati kontroll Ensembl veebilehel (versioon 84) (Yates *et al.*, 2016).

⁹ cancer.sanger.ac.uk

Tabel 1. Amplikonipõhiseks teise põlvkonna sekveneerimiseks valitud geenide eksonid ja lookusspetsiifiliste praimerite järjestused.

Geen	Ekson	Lookusspetsiifilise praimeride järjekord (5'-3')	
BRAF	11	<i>forward</i>	TGGGCAGATTACAGTGGGACA
		<i>reverse</i>	CCACATTACATACTTACCATGCCAC
	15	<i>forward</i>	ATTTCTTCATGAAGACCTCACAGTA
		<i>reverse</i>	ACTGTTCAAACCTGATGGGACC
EGFR	18	<i>forward</i>	TCTCTTGAGGATCTTGAAGGAAAC
		<i>reverse</i>	TGTGCCAGGGACCTTACCTTATA
	19	<i>forward</i>	TGAGAAAGTTAAAATTCCCGTCGC
		<i>reverse</i>	CCCCACACAGCAAAGCAGAA
	20	<i>forward</i>	TCCAGGAAGCCTACGTGATG
		<i>reverse</i>	AGGCAGCCGAAGGGCATGA
	21	<i>forward</i>	AACACCGCAGCATGTCAAGA
		<i>reverse</i>	CACCTCCTTACTTTGCCTCCTTC
ERBB2	19	<i>forward</i>	TCCCTGATGGGGAGAATGTGA
		<i>reverse</i>	GGTGGAGGGGCTTACGTCT
	20	<i>forward</i>	TACCCTTGTCCTCCAGGAAGCAT
		<i>reverse</i>	GCACCGTGGATGTCAGGCAG
KRAS	2	<i>forward</i>	AGGCCTGCTGAAAATGACTGAA
		<i>reverse</i>	TGAATTAGCTGTATCGTCAAGGCA
	3	<i>forward</i>	CTTGGATATTCTCGACACAGCA
		<i>reverse</i>	AAAGAAAGCCCTCCCCAGTC
PIK3CA	10	<i>forward</i>	CAGCTCAAAGCAATTTCTACACGA
		<i>reverse</i>	TCCATTTTAGCACTTACCTGTGACT
	21	<i>forward</i>	AGCCTTAGATAAAACTGAGCAAGAGG
		<i>reverse</i>	TTTGTTGTCCAGCCACCAT

Disainitud praimeritele lisati 5' otsa Illumina® üleulatuvad adapterid (*forward overhang*: 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[lookusspetsiifiline järjekord] ja *reverse overhang*: 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-[lookus-

spetsiifiline järjestus]) ning saadud praimerjärjestuste omavahelist sobivust juuksenõelastruktuuride ja dimeeride suhtes kontrolliti OligoAnalyzer[®] 3.1¹⁰ (Integrated DNA Technologies, Inc., USA) ja MultiPLX 2.1 (Kaplinski *et al.*, 2005) programmiga. Viimati nimetatud veebitööriistaga jagati praimerid kahte, kolme ja nelja gruppi, arvestades seondumistugevusi, produkti pikkust, sulamistemperatuuri ning seondumistugevusi omavahel ja PCR produktidega.

Praimerite efektiivsust kontrolliti PCR-ga nii üksikult kui ka grupeeritult inimese genoomse DNA (Human Genomic DNA, Roche Holding AG, Šveits) ja tervete isikute rakuvaba DNA segu peal. Reaktsioonisegu kogumahuks oli 20 µl, mis sisaldas lõppkontsentratsioonis 1x HOT FIREPol[®] GC Master Mix-i, 2,5 mM MgCl₂, 10 pmol praimereid ja 4 ng cfDNA-d või 45 ng gDNA-d. Amplifikatsioon viidi läbi Arktik[™] termotsükleris järgmistel tingimustel:

polümeraasi aktivatsioon	95 °C 12 min	
denaturatsioon	95 °C 30 sek	} cfDNA puhul 45, gDNA puhul 35 tsükli
praimerite seondumine	58 °C 30 sek	
elongatsioon	72 °C 30 sek	
lõplik elongatsioon	72 °C 5 min	

programmi lõpp, temperatuuri hoidmine 4 °C.

PCR-i produktide olemasolu kontrolliti 3% agarosgeelil (3% agaros, 0,5X TBE (Tris-boorhape-EDTA)).

Sekveneerimisraamatukogu valmistamine

Teise põlvkonna sekveneerimise jaoks tehti raamatukogud cfDNA-st, FFPE DNA-st ja erinevate mutatsiooni protsentidega *spike-in*-dest, kokku 50 proovi (tabel 2).

¹⁰ <http://www.idtdna.com/Scitools>

Tabel 2. Proovid, millest tehti raamatukogude segu MiSeq® platvormil (Illumina, Inc.) sekveneerimiseks. Multipleks reaktsioone – erinevate praimerigruppide arv; tsüklite arv – multipleks amplikon PCR-il kasutatud tsüklite arv.

Proov	Multipleks reaktsioone	Tsüklite arv
cfDNA 1. replikaat	7	45
cfDNA 2. replikaat	7	45
cfDNA	7	35
FFPE DNA	7	35
0,2% deletsiooni ja punktmutatsiooni <i>spike-in</i>	4	35
0,4% deletsiooni ja punktmutatsiooni <i>spike-in</i>	4	35
1% deletsiooni ja punktmutatsiooni <i>spike-in</i>	4	35
2% deletsiooni ja punktmutatsiooni <i>spike-in</i>	4	35
4% deletsiooni ja punktmutatsiooni <i>spike-in</i>	4	35
50% deletsiooni <i>spike-in</i>	1	35
100% deletsiooni kloon	1	35

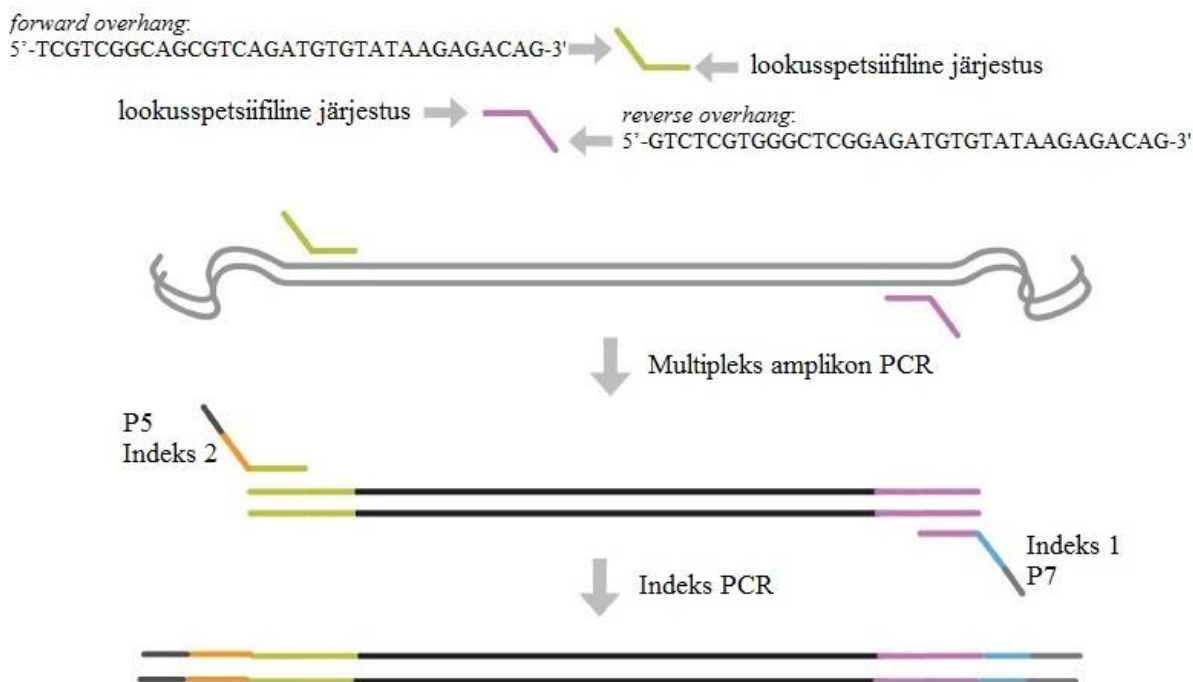
Raamatukogu loomiseks teostati esmalt multipleks amplikon PCR. Tulemusena saadi produktid, mille 5' otstes on üleulatuvad adapterid. Reaktsioonisegu kogumahuks oli 20 µl, mis sisaldas lõppkontsentratsioonis 1x HOT FIREPol® GC Master Mix-i, 2,5 mM MgCl₂, 10 pmol praimereid ja 1,5 ng DNA-d. Amplifikatsioon viidi läbi Arktik™ termotsükleris järgmistel tingimustel:

polümeraasi aktivatsioon	95 °C 12 min	
denaturatsioon	95 °C 30 sek	} 35 või 45 tsüklit
praimerite seondumine	58 °C 30 sek	
elongatsioon	72 °C 30 sek	
lõplik elongatsioon	72 °C 5 min	

programmi lõpp, temperatuuri hoidmine 4 °C.

Järgnevalt kasutati Tartu Ülikooli Eesti Geenivaramu tuumiklabori protokoll PCR produktide puhastamiseks Agencourt AMPure XP (Beckman Coulter, Inc., USA) magnetkerakestega

(lisa 3 ja 3.1) ja indeks PCR-i läbiviimiseks, mille tulemusena liideti ampliconidele Nextera[®] XT indeksjärjestused (Illumina, Inc., USA) (joonis 6).



Joonis 6. Ampliconide märkimine Illumina[®] üleulatuivate adapterite ja indeksitega (Illumina, Inc.).

Indeks PCR-i reaktsioonisegu kogumahuks oli 50 µl, mis sisaldas lõppkontsentratsioon 1x KAPA SYBR[®] FAST qPCR Kit Master Mix Universal-i (Kapa Biosystems, Inc., USA), 5 µl Nextera[®] XT i7 indeks praimer 1 (Nextera[®] XT v2 Index Kit, Illumina, Inc.), 5 µl Nextera[®] XT i5 indeks praimer 2 (Nextera[®] XT v2 Index Kit, Illumina, Inc.) ja 5 µl puhastatud amplicon PCR-i produkte. Indeksjärjestuste liitmine viidi läbi Arktik[™] termotsükleris järgmistel tingimustel:

polümeraasi aktivatsioon	95 °C 10 min	
denaturatsioon	95 °C 30 sek	} 8 tsükli
praimerite seondumine	55 °C 30 sek	
elongatsioon	72 °C 30 sek	
lõplik elongatsioon	72 °C 5 min	

programmi lõpp, temperatuuri hoidmine 4 °C.

Indeksitega märgitud PCR-i produktid puhastati samuti lisas 3 toodud viisil. Valmis raamatukogude kontsentratsioon mõõdeti Qubit[®] 2.0 fluoromeetriga, kasutades Qubit[®]

dsDNA HS Assay kitti. Amplikonide pikkust mõõdeti TapeStation 2200 instrumendil, kasutades D1000 ScreenTape-i[®]. Leidmaks raamatukogude molaarset kontsentratsiooni (mol/l), kasutati valemit:

$$\frac{\text{kontsentratsioon ng/}\mu\text{l}}{660 \text{ g/mol} \times \text{keskmine pikkus}} \times 10^6 = \text{lahuse nanomolaarne (nM) kontsentratsioon.}$$

Arvutatud nanomolaarse kontsentratsiooni põhjal lahjendati raamatukogud esmalt 10 mM TRIS pH 8,5 puhvris 10 nM-ks. Saadud lahjenduste kontsentratsiooni kontrolliti Qubit[®] 2.0 fluoromeetriga, kasutades Qubit[®] dsDNA HS Assay kitti ning lahjendati 10 mM TRIS pH 8,5 puhvris edasi 4 nM-ks. Kõik 4 nM raamatukogud segati omavahel võrdses koguses kokku.

Raamatukogu kvantifitseerimine ja MiSeq sekvenaatorisse laadimine

Leidmaks reagentide kassetile (*reagent cartridge*) laadimiseks vajalikku raamatukogude segu pikomolaarset (pM) kontsentratsiooni, tehti võrdlev qPCR varem sekveneeritud raamatukoguga. Reaktsioonisegu kogumahuks oli 10 µl, mis sisaldas lõppkontsentratsioonis 1x KAPA SYBR[®] FAST qPCR Kit Master Mix Universal-i koos praimeritega (Kapa Biosystems, Inc., USA) ja KAPA standardeid (Kapa Biosystems, Inc.). DNA paljundamiseks qPCR-i käigus kasutati raamatukogude 2,5 pM ja 5 pM lahjendust. Amplifikatsioon viidi läbi 7900HT Fast Real-Time PCR System termotsükleris järgmistel tingimustel:

polümeraasi aktivatsioon	95 °C 5 min	
denaturatsioon	95 °C 30 sek	} 35 tsüklit.
praimerite seondumine ja elongatsioon	60 °C 45 sek	

Tulemuste vaatamiseks kasutati SDS v2.4 tarkvara (Applied Biosystems, USA). Võttes arvesse eelmisel korral laetud raamatukogu kontsentratsiooni, saadud klastrite tihedust (K/mm²) ja qPCR-i tulemusena saadud sekveneeritava raamatukogu kontsentratsiooni, arvutati soovitud 950 K/mm² tiheduse saavutamiseks vajalik raamatukogu pikomolaarne kontsentratsioon. Seejärel koostati Illumina[®] Experiment Manager (IEM) v1.11 tarkvara (Illumina, Inc., USA) kasutades *SampleSheet* ja *Manifest* faili. Esimest neist on vaja MiSeq[®] sekvenaatori (Illumina, Inc.) käivitamiseks ja hilisemaks analüüsiks ning teist MiSeq Reporter analüüsitarkvara (Illumina, Inc.) käivitamiseks. *SampleSheet* sisaldab informatsiooni proovide kohta, näiteks proovi nime, igale proovile vastavat indeksjärjestust ja viidet *Manifest* failile. *Manifest* fail sisaldab informatsiooni amplikonide kohta, näiteks nende asukoht kromosoomil ja praimerite pikkused. Enne raamatukogude segu laadimist reagentide kassetile, teostati segu denatureerimine 0,2 N NaOH-ga ja lahjendamine soovitud kontsentratsioonini ning 0,5%

PhiX kontrolliga kokkusegamine vastavalt MiSeq[®] Reagent Kit v2 protokollile (Illumina, Inc.). Raamatukogu sekveneeriti MiSeq[®] sekvenaatoriga (Illumina, Inc.), 2 x 116 bp järjestustega (*paired-end*).

Sekveneerimistulemuste analüüs

Illumina[®] Sequencing Analysis Viewer v1.10.2 (SAV) tarkvara (Illumina, Inc.) kasutades kontrolliti sekveneerimise kvaliteedi parameetreid, nagu näiteks Q30, klastrite tihedust ja järjestuste arvu. Q30 kvaliteediskoori korral on sekveneerimisvea sagedus 1:1000 ehk 99,9% tõenäosusega on ahelasse lülitunud õiged nukleotiidid.

MiSeq[®] sekvenaatorist saadi järjestuste toorandmed .fastq failiformaadis. Indelite ja ühenukleotiidiliste asenduste leidmiseks kasutati kolme erinevat meetodit. Esiteks rakendati Amplicon Indel Hunter v1.0.0 (aiHunter) programmi (Kadri *et al.*, 2015) parameetritega --cushion 3 --maf 0.01, et detekteerida indeleid, mis on pikemad kui 3 bp ja mille MAF on suurem kui 1%.

Teiseks kasutati sekveneeritud FASTQ järjestuste joondamiseks GRCh37 inimese referentsgenoomile BWA-MEM algoritmi (Li, 2013), mille väljundiks on .bam failid. Saadud .bam faile protsessiti edasi Genome Analysis Toolkit v3.5 (GATK) (McKenna *et al.*, 2010) HaplotypeCaller algoritmiga, mille väljundiks on variandid .vcf laiendiga failis. Leidudeks loeti ainult sellised ühenukleotiidilised asendused, millel oli kõrge kvaliteet (QD > 5), mille puuduvate väärtuse osakaal oli <5% ning mis ei moodustanud valepositiivseid klastreid (3 ühenukleotiidilisist asendust 10 bp sees). Seejärel annoteeriti tulemused Variant Effect Predictor (VEP) (McLaren *et al.*, 2010) tarkvaraga, määraes geeni ja transkripti, kus variant asub, variandi tähtsuse ja asukoha valgu järjestusel. Annotatsioon sisaldas ka päringuid teistest andmebaasidest nagu OMIM, HGMD, ClinVar ja Cosmic.

Lisaks visualiseeriti .bam laiendiga faile Integrative Genomics Viewer (IGV) tarkvaraga (Robinson *et al.*, 2011), seades alleeli fraktsiooni läveks 1%. Programmi abil otsiti käsitsi ühenukleotiidilisi asendusi ning teadaolevaid deletsioone.

Uuritavad

27.05.2016 seisuga on „Tsirkuleeriv DNA kopsuvähi biomarkerina“ uuringusse kaasatud 80 mitteväikerakulise kopsuvähiga patsienti. Nendest esimesed 70 isikut moodustavad antud magistritöö valimi (tabel 3). Uuringus osalemiseks andsid kopsuvähiga patsiendid informeeritud nõusoleku. Vereproove võttis SA Põhja-Eesti Regionaalhaigla (PERH) uuringuõde ning kliiniliste andmete kogumisel osalesid onkoloog ja projekti käigus kaastööd tegev TÜ arstiteaduse üliõpilane. Kõiki koematerjale ja kliinilisi andmeid säilitatakse ning kasutatakse kodeerituna.

Tabel 3. Magistritöösse kaasatud mitteväikerakulise kopsuvähiga patsientide kokkuvõttev tabel. Igale uuringu koodile vastab üks isik. Sugu: N – naine, M – mees. Histoloogiline leid: AK – adenokartsinoom, LR – lamerakuline kartsinoom, SR – suurerakuline kartsinoom, TK – täpsustamata kartsinoom, TMVR – täpsustamata mitteväikerakuline kartsinoom, VR - väikerakuline kartsinoom. TNM – vt lk 14. St – staadium: III või IV (kaugmetastaasid). TTF-1, p63, CK5 ja CK6 – immuunohistokeemilised markerid. EGFRm diag lab – diagnostikalaboris saadud EGFR geeni mutatsioonianalüüsi tulemus. Vähir osakaal – vähirakkude osakaal FFPE koematerjalis.

Kood	Sugu	Vanus	Esmane diagn	Histol leid	T	N	M	St	TTF-1	p63	CK5	CK6	EGFRm diag lab	Vähir osakaal
kA02	N	85	jah	AK			1a	IV					pos	50%
kA03	M	58	jah	LR	2	2	1	IV	pos					
kA04	M	72	jah	AK	3	2	0	IIIA	pos	neg				
kA05	M	48	jah	AK		0	1	IV	pos		neg	neg	neg	30%
kA06	N	66	jah	VR	3	2		IIIA						
kA07	N	49	jah	AK		3	1b	IV	pos				neg	
kA08	N	64	jah	AK	2a	1	0	IIA					neg	50%
kA09	M	64	jah	AK	2	3	1	IV	neg		pos	pos		
kA10	N	75	jah	AK	2b		1a	IV	pos	neg			pos	30%
kA11	M	54	jah	TK	2a	2	1a	IV						
kA12	M	63	jah	TK	4	3	1a	IV	neg	pos	pos	pos	neg	75%
kA13	M	61	jah	TK	2a	2	1b	IV						
kA14	M	81	jah	AK	3	2		IIIA	pos	neg			neg	90%
kA15	M	46	retsidiiv	AK		2	1a	IV	pos		neg	neg	neg	30%
kA16	N	80	jah	AK	2a	2	1a	IV	pos				neg	30%
kA17	M	67	jah	TK	2b	2		IIIA						
kA18	N	57	jah	VR	4	3	1b	IV	neg		neg	neg		
kA19	N	75	jah	AK	4	2	1a	IV					pos	50%
kA20	M	59	jah	AK	3		1	IV					pos	30%
kA21	M	76	jah	LR	4	2	1	IV						
kA22	N	75	jah	AK	2		1a	IV	pos				neg	10%
kA23	M	66	jah	TK	4	2	1b	IV						
kA24	N	57	jah	AK	1b	2	0	IIIA					pos	80%
kA25	M	68	jah	AK	1b	3	1	IV	pos				neg	90%

kA26	N	81	retsidiiv	AK	2		1	IV					pos	90%
kA27	N	70	jah	AK	2a		1a	IV	pos				neg	
kA28	N	54	jah	AK	3	3	1	IV	pos	neg			pos	25%
kA29	M	74	retsidiiv	AK	2a	2	1a	IV	pos					
kA30	M	50	jah	AK	2a	3	1	IV					neg	15%
kA31	M	82	jah	AK	1b	2	0	IIIA	pos					
kA32	M	60	jah	TK	2a	2	0	IIIA	pos				neg	80%
kA33	N	66	retsidiiv	AK	2	3	0	IIIB	pos				pos	
kA34	N	74	jah	AK		2	1a	IV	pos		neg	neg	neg	20%
kA35	N	49	jah	AK	3	2	1	IV	neg					
kA36	M	71	retsidiiv	AK	2a	2	0	IIIA	pos		neg	neg	neg	80%
kA37	M	31	jah	AK	4	1	1	IV	neg	neg	neg	neg	neg	30%
kA38	N	59	jah	AK	2	0	1b	IVB	pos				neg	10%
kA39	M	73	jah	AK	2a	2	1a	IV	pos	pos	neg	neg	neg	50%
kA40	N	55	jah	AK	2a	2	1b	IV	pos	neg	neg	neg	neg	30%
kA41	N	56	jah	AK	4	3	1	IV	pos		neg	neg	neg	70%
kA42	M	77	jah	AK	2	2	1	IVB	pos	neg	neg	neg		
kA43	M	69	jah	AK	4	3		IIIB	pos	neg				
kA44	M	72	retsidiiv	AK	3	2		IIIA	pos				neg	50%
kA45	N	71	retsidiiv	AK	2a	0	0	IB	pos		neg	neg	neg	15%-70%
kA46	M	68	retsidiiv	AK	1a	0	1	IV	pos				neg	90%
kA47	N	70	jah	AK	2a	2	1	IV	pos	neg	neg	neg	pos	3%
kA48	N	70	jah	AK	2b		1b	IV		neg			pos	100%
kA49	N	61	jah	AK	3	2		IIIA	pos	neg	neg	neg	neg	30%
kA50	N	53	jah	AK	2	2	1b	IV					neg	30%
kA51	M	69	jah	AK	1a	2	1b	IV	pos	pos	neg	neg	neg	40%
kA52	M	62	jah	AK	2b	3	1b	IV	pos				neg	20%
kA53	M	66	jah	AK	4	0	1a	IV					neg	10%
kA54	M	79	jah	AK	2a	3	1b	IV	pos				neg	
kA55	M	67	jah	AK	4	3		IIIB	pos				neg	50%
kA56	N	67	jah	SR	3		0	IIB	pos	neg	neg	neg		
kA57	M	72	jah	AK	2a	3		IIIB	neg		pos	pos	neg	30%
kA58	M	64	jah	AK	2	3	1	IV	pos		neg	neg	neg	80%
kA59	N	67	jah	AK	4	2	1a	IV					pos	15%
kA60	N	56	jah	TK	1b	3	1b	IV						
kA61	M	74	retsidiiv	AK	3	2	0	IIIA	pos	neg			neg	50%
kA62	N	55	jah	TK	3	3		III	neg	pos	neg	neg	pos	30%
kA63	N	80	jah	AK	4	2		IIIB	pos	pos			pos	30%
kA64	M	63	jah	AK	4b	1	1	IV	pos	neg	neg	neg	neg	50%
kA65	N	70	jah	AK	2a	2	0	IIIA		neg	neg	neg	neg	50%
kA66	M	68	jah	TMVR	4	2	1b	IV	neg	neg				
kA67	N	74	jah	AK		2	1	IV	pos	neg			pos	80%
kA68	N	74	jah	AK	1b	0	0	IA	pos				neg	50%
kA69	M	71	jah	AK	2b	2	1b	IV	neg	neg			neg	

kA70	M	59	jah	AK		2	1a	IV	pos	pos	neg	neg	neg	90%
kA71	M	53	jah	AK	2b	3	1	IV			neg	neg	neg	30%

Uuritavate keskmine vanus uuringusse kaasamise hetkel oli 66 aastat (vahemikus 31...85), nendest 46% olid naised. 70 isikust 56 (80%) oli histoloogiliseks leiuks adenokartsinoom. Kõige enam oli IV kliinilise staadiumiga (kopsuvähi kaugmetastaasidega) isikud (69%).

Kõigilt 70 isikult oli kogutud veeniveri ning 4 isiku (kA20, kA28, kA34 ja kA51) puhul oli kasvaja progresseerumise järel kogutud ka korduv materjal. 33 uuritava (47%) puhul oli võimalik saada SA PERH patoloogiakeskusest sama isiku FFPE koematerjal. Uuringusse saatmiseks oli koematerjal lõigatud 10 µm paksusteks koelõikudeks. SA PERH patoloogiakeskusest saadeti ainult selliseid koelõike, kus oli eelneval tuvastatud kasvajamaterjali esinemine.

Igalt patsiendilt koguti 2 Cell-Free DNA BCT[®] CE 10 ml (Streck, Inc., USA) katsutit veeniverd. Lisaks vereplasmale koguti vajadusel iduliini geneetiliste variantide määramiseks ka leukotsüütide rikas kiht (*buffy coat*), mida säilitati samadel tingimustel.

Uuritavate rakuvaaba ja koelõigu DNA

Patsientide plasmast rakuvaaba DNA ja FFPE koelõigust genoomse DNA eraldamiseks kasutati vastavalt QIAamp[®] Circulating Nucleic Acid ja GeneRead[™] DNA FFPE kitti, tootjate poolt etteantud instruktsioonide kohaselt, kuid kasutades soojendatud elueerimispuhvrit ning kaks korda elueerimist. Eraldatud DNA kontsentratsiooni mõõdeti Qubit[®] 2.0 fluoromeetriga kasutades cfDNA puhul Qubit[®] dsDNA HS Assay ja FFPE DNA puhul Qubit[®] dsDNA BR Assay kitti.

Ravimseoseliste mutatsioonide testimine

Uuritavate cfDNA-d ja FFPE DNA-d kontrolliti ravimseoseliste mutatsioonide suhtes eelnevalt väljatöötatud fragmentanalüüsi ja teise põlvkonna sekveneerimise meetoditega, selle erinevusega, et fragmentanalüüsile eelnev PCR amplifikatsioon viidi läbi Arktik[™] termotsükleris ning NGS analüüsile eelnev amplikon PCR-i reaktsioonisegu kogumaht oli 25 µl. Amplikonipõhiseks teise põlvkonna sekveneerimiseks koostati raamatukogud kõigist 107-st uuritavatelt kogutud proovidest (70 cfDNA esmast materjali, 4 cfDNA korduvat progresseerunud materjali, 33 FFPE DNA materjali) ning tervete kontrollide cfDNA segust, 1% c.2235_2249del15 deletsiooni ja 1% c.2573T>G punktmutatsiooni *spike-in*-dest.

Sekveneerimisel piisava katvuse saamiseks jagati proovid kaheks raamatukoguks, mida sekveneeriti eraldi.

Mutatsioonileidude võrdlemine

Mutatsioonileidude võrdlemiseks kasutati esmase etalonina sertifitseeritud diagnostikalaboris saadud EGFR analüüsitulemusi. Tulemuste hindamisel tehti esmalt oletus, et diagnostikalabori tulemus on absoluutselt tõene. Diagnostikalaboris oli testitud ainult FFPE koematerjali EGFR geeni kahe peamise mutatsiooni suhtes, seejuures pole ravi määramise seisukohalt üldjuhul oluline teha vahet 19. eksoni deletsioonidel ja c.2573T>G punktmutatsioonil. Sellise lähenemisega oli võimalik kontrollida vaid 28 uuritava EGFR geeni mutatsioonileide. Diagnostikalaboris saadud tulemuste põhjal jagati välja töötatud meetoditega saadud tulemused nelja klassi: õige positiivsed (*true positive*, TP) ja negatiivsed (*true negative*, TN) ning valepositiivsed (*false positive*, FP) ja -negatiivsed (*false negative*, FN). Nende põhjal arvutati fragmentanalüüsi ning teise põlvkonna sekveneerimise meetodite tundlikkus, spetsiifilisus, positiivne ennustusväärtus (*positive predictive value*, PPV) ja kokkulangevus (*concordance*) diagnostikalabori tulemustega. NGS tulemusi analüüsiti kolme erineva meetodikaga: käsitsi järjestuste läbivaatamine IGV-s (manuaalne) ja programmidega aiHunter või GATK (tarkvarapõhine). Uuritavatel, kelle puhul polnud FFPE koelõike võimalik saada (37/70) või puudusid diagnostikalabori tulemused (5/33), määrati mutatsioonileid õigeks positiivseks, kui vähemalt kaks analüüsimeetodit kolmest (fragmentanalüüs, NGS manuaalne ja tarkvarapõhine) tuvastasid mutatsiooni esinemist.

Tulemused

Käesoleva magistritöö eksperimentaalse osa esimeses pooles töötati välja fragmentanalüüsil ja teise põlvkonna sekveneerimisel põhinevad meetodid cfDNA-st ja FFPE DNA-st ravim tundlikkust mõjutavate mutatsioonide tuvastamiseks. Töö teises pooles eraldati mitteväikerakulise kopsuvähiga patsientidelt kogutud verest cfDNA ja FFPE koelõigust gDNA ning testiti neid huvipakkuvate mutatsioonide suhtes eelnevalt väljatöötatud fragmentanalüüsi ja NGS meetoditega.

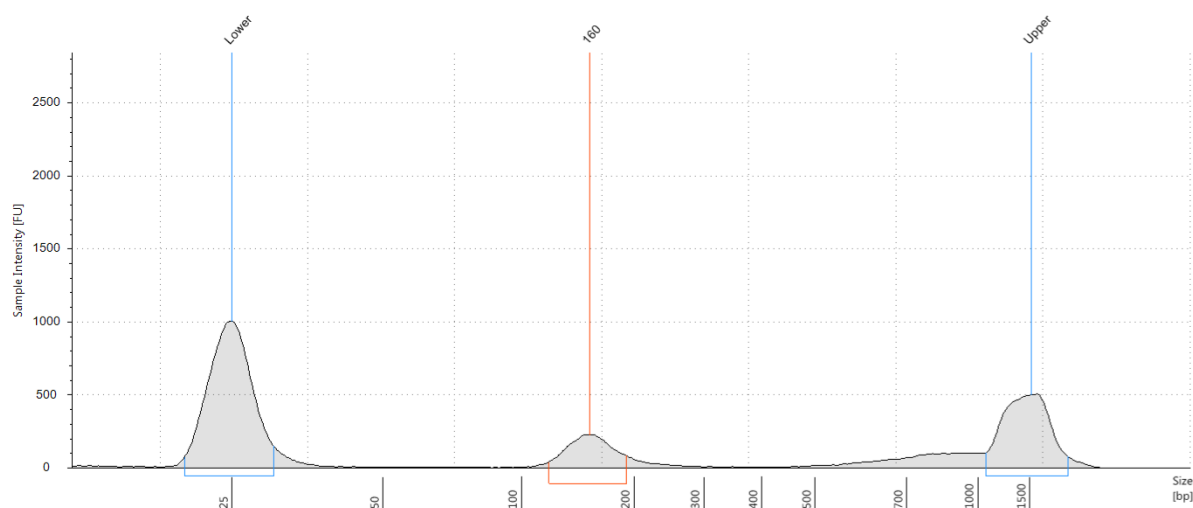
Analüüsimetoodikate väljatöötamine ja optimeerimine

Mitteväikerakulise kopsuvähiga patsientide rakuvaba ning koelõigu genoomse DNA analüüsimiseks somaatiliste mutatsioonide suhtes tuli esmalt välja töötada meetodid fragmentanalüüsile ja teise põlvkonna sekveneerimisele eelnevaks huvipakkuvate regioonide paljundamiseks ning sekveneerimistulemuste analüüsiks.

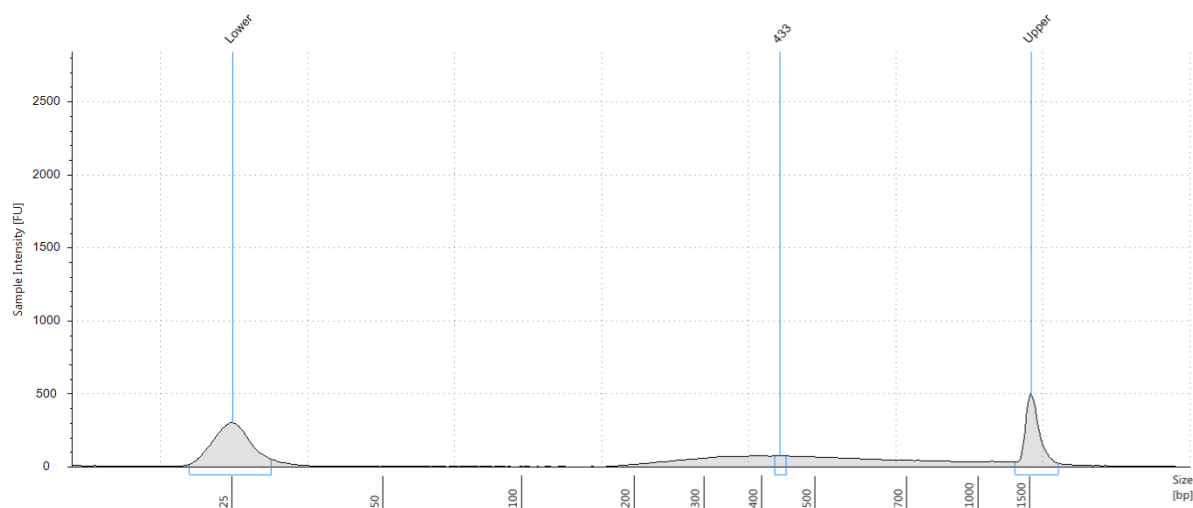
Rakuvaba ja FFPE DNA eraldamine ning kvaliteedi kontroll

Tervetelt vabatahtlikelt kogutud plasmast rakuvaba DNA eraldamiseks kasutati kahte kitti (NucleoSpin® Plasma XS ja QIAamp® CNA), mille puhul võrreldi DNA saagikust kontsentratsioonide alusel. NucleoSpin® Plasma XS kitiga eraldatud DNA kontsentratsiooni polnud enamikul juhtudel võimalik Qubit® 2.0 fluoromeetriga mõõta, mis tähendab, et proovide kontsentratsioon oli alla 0,01 ng/μl. Paari proovi puhul oli siiski võimalik kontsentratsiooni mõõta ning nendel juhtudel saadi DNA keskmiseks saagikuseks 1,77 ng (2,46 ng cfDNA/ml plasma kohta). QIAamp® CNA kitiga eraldatud DNA madalaim kontsentratsioon oli võrdluseks 0,11 ng/μl ning keskmine saagikus oli 12,39 ng (2,54 ng cfDNA/ml plasma kohta). Saadud tulemuste ja kirjanduse (Devonshire *et al.*, 2014, Mauger *et al.*, 2015, Page *et al.*, 2013) põhjal otsustati edaspidi rakuvaba DNA eraldamiseks kasutada QIAamp® Circulating Nucleic Acid kitti. FFPE DNA eraldamiseks GeneRead™ DNA FFPE kitti kasutades saadi DNA saagikuseks 4,2 μg.

TapeStation 2200 instrumendiga mõõtes jäi rakuvaba DNA fragmentide pikkus vahemikku 125...200 aluspaari (*basepair*, bp) (joonis 7) ning FFPE koelõigust eraldatud genoomse DNA puhul jäi fragmentide pikkus vahemikku 200...1000 bp (joonis 8).



Joonis 7. Rakuvaba DNA fragmentide pikkus mõõdetuna TapeStation 2200 instrumendil. X-teljel fragmentide pikkus aluspaarides (bp), y-teljel proovi intensiivsus fluorestsents ühikutes (FU). *Lower* – alumine pikkusmarker, *Upper* – ülemine pikkusmarker. Antud proovi puhul jääb cfDNA fragmentide pikkus vahemikku 125...200 bp.



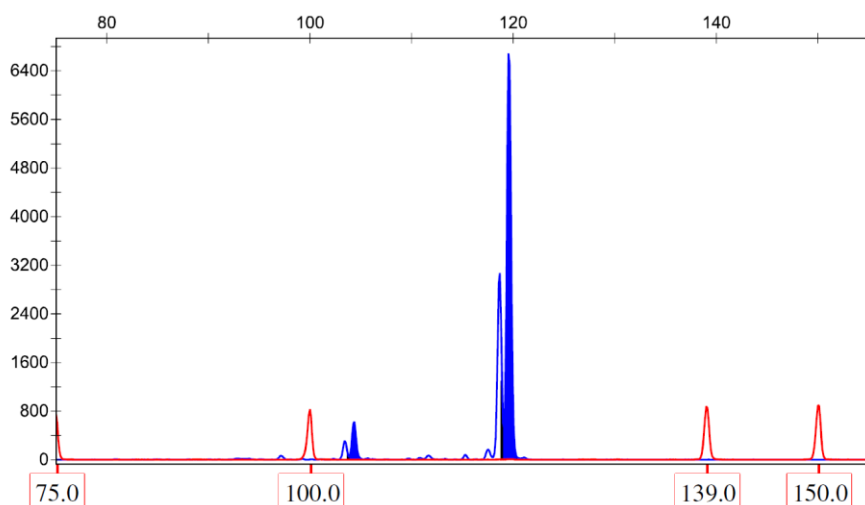
Joonis 8. FFPE koelõigust eraldatud genoomse DNA fragmentide pikkus mõõdetuna TapeStation 2200 instrumendil. X-teljel fragmentide pikkus aluspaarides (bp), y-teljel proovi intensiivsus fluorestsents ühikutes (FU). *Lower* – alumine pikkusmarker, *Upper* – ülemine pikkusmarker. Antud proovi puhul jääb FFPE DNA fragmentide pikkus vahemikku 200...1000 bp.

Fragmentanalüüs

Fragmentanalüüsi meetodi tundlikkuse ja spetsiifilisuse määrab sellele eelnevalt tehtud huvipakkuvate regioonide amplifitseerimine PCR-i käigus. EGFR geeni c.2235_2249del15 ja c.2573T>G mutatsiooni võimalikult spetsiifiliseks ja tundlikuks tuvastamiseks prooviti nimetatud mutatsiooni kloonide *spike-in* segusid amplifitseerida erinevate PCR-i reaktsioonisegude ja –tingimustega.

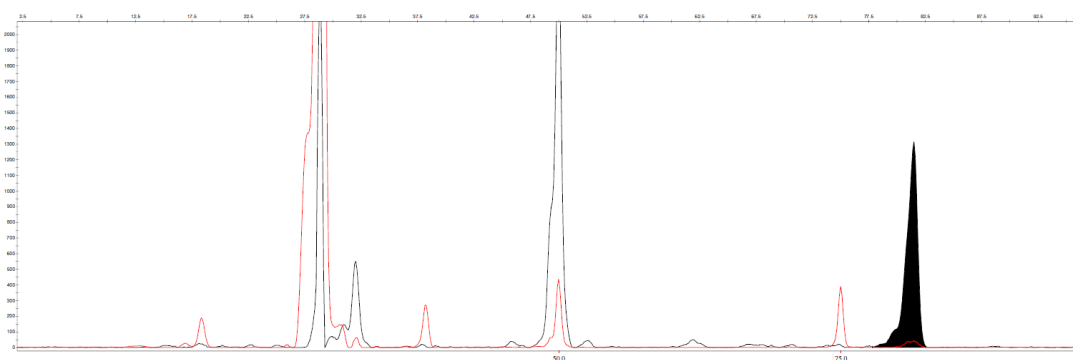
Kasutades leheküljel 23 välja toodud protokoll, oli EGFR geeni 19. eksoni c.2235_2249del15 deletsiooni tuvastamise tundlikkus algselt 4% (joonis 9). Seda järelitati sellest, et madalamate

protsentidega (0,2%, 0,5%, 1%, 2%) *spike-in* segude puhul oli elektroferogrammil näha ainult metsiktüüpi piik. PCR-i käigus produktide paljundamise suhtelist saagikust mõõdeti elektroferogrammi piigi kõrguse järgi.



Joonis 9. EGFR geeni 19. eksoni c.2235_2249del15 deletsiooni 4% *spike-in* segu fragmentanalüüsi reaktsiooni tulemusena saadud elektroferogramm. X-teljel fragmentide pikkus aluspaarides (bp), y-teljel suhteline fluorestsents ühik (*relative fluorescence unit*, RFU). Punased piigid – pikkusmarkerid, sinised täidetud piigid – paljundatud produktid (106 bp mutantne, 121 bp metsiktüüpi).

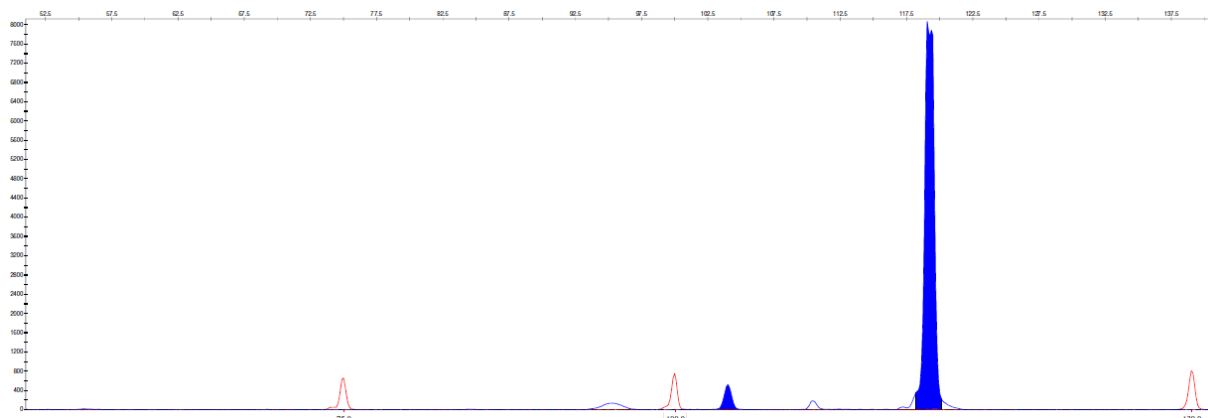
Kasutades leheküljel 23 välja toodud protokoll, oli EGFR geeni 21. eksoni c.2573T>G punktmutatsiooni tuvastamise tundlikkus algselt rohkem kui 8% (joonis 10). Seda järelitati sellest, et kasutatud reaktsioonitingimustel polnud ühegi kasutatud *spike-in* (0,2%, 0,5%, 1%, 2%, 4%, 8%) segu puhul mutantset produkti amplifitseerunud.



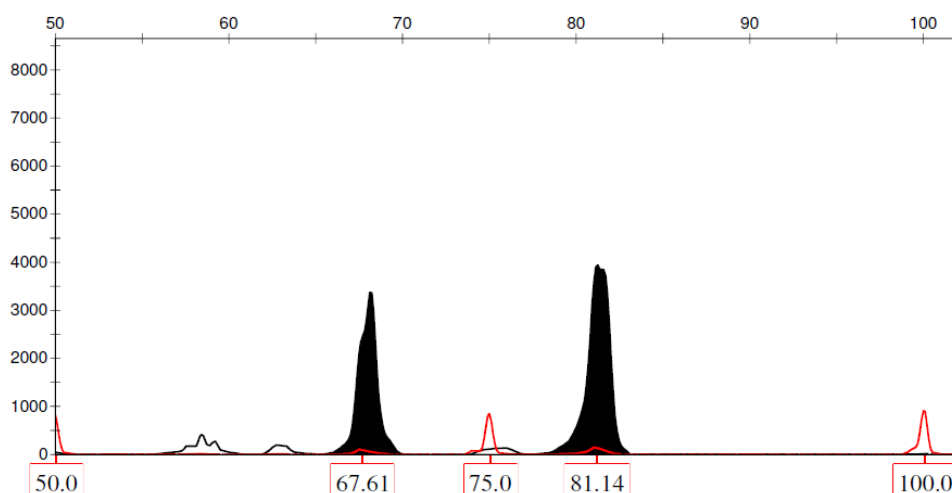
Joonis 10. EGFR geeni 21. eksoni c.2573T>G punktmutatsiooni 8% *spike-in* segu fragmentanalüüsi reaktsiooni tulemusena saadud elektroferogramm. X-teljel fragmentide pikkus aluspaarides (bp), y-teljel suhteline fluorestsents ühik (*relative fluorescence unit*, RFU). Punased piigid – pikkusmarkerid, must täidetud piik – paljundatud 80 bp metsiktüüpi produkt.

Pärast reaktsioonisegude ja –tingimuste modifitseerimist saadi lõplikuks fragmentanalüüsi meetodil EGFR geeni c.2235_2249del15 ja c.2573T>G mutatsiooni tuvastamise

tundlikkuseks 0,5% (joonis 11 ja 12). Selleks kasutati EGFR geeni 19. eksoni puhul leheküljel 24 ja 21. eksoni puhul leheküljel 25 välja toodud protokoli.



Joonis 11. EGFR geeni 19. eksoni c.2235_2249del15 deletsiooni 0,5% *spike-in* segu fragmentanalüüsi reaktsiooni tulemusena saadud elektroferogramm. X-teljel fragmentide pikkus aluspaarides (bp), y-teljel suhteline fluorestsents ühik (*relative fluorescence unit*, RFU). Punased piigid – pikkusmarkerid, sinised täidetud piigid – paljundatud produktid (106 bp mutantne, 121 bp metsiktüüpi).



Joonis 12. EGFR geeni 21. eksoni c.2573T>G punktmutatsiooni 0,5% *spike-in* segu fragmentanalüüsi reaktsiooni tulemusena saadud elektroferogramm. X-teljel fragmentide pikkus aluspaarides (bp), y-teljel suhteline fluorestsents ühik (*relative fluorescence unit*, RFU). Punased piigid – pikkusmarkerid, mustad täidetud piigid – paljundatud produktid (67 bp mutantne, 80 bp metsiktüüpi).

Väljatöötatud PCR-i reaktsioonisegu ja –tingimusi kasutati ka uuritavate cfDNA ja FFPE DNA analüüsimiseks EGFR geeni 19. eksoni deletsioonide ja 21. eksoni c.2573T>G punktmutatsiooni suhtes.

TruSight® Tumor 26 paneel

Eraldatud cfDNA ja FFPE DNA fragmentide pikkuse mõõtmisel saadud tulemused (joonised 7 ja 8) andsid alust arvata, et nimetatud proovid võiksid sobida Illumina® TruSight® Tumor 26

paneeli sekveneerimisraamatukogu tegemiseks, kuna nimetatud kitis on paljundatavate amplikonide pikkus 165-195 bp¹¹.

Illumina® TruSight® Tumor 26 paneeli puhul hinnati esmalt eraldatud DNA sobivust, võttes aluseks qPCR-i tulemusena saadud ΔCq väärtused. Iga proovi ΔCq väärtus võrdub proovi keskmine Cq (*quantitation cycle*, tsükkel, alates millest detekteeritakse fluorestsentsi) väärtus miinus QCT (*quality control template*) keskmine Cq väärtus. Illumina® ei soovita kasutada proove, mille ΔCq väärtus on suurem kui 4. Tehtud kvaliteedikontrollil saadi rakuvaba DNA ΔCq väärtuseks 100-kordse lahjenduse puhul 10,04. Seejärel otsustati QCT-ga võrrelda ka 10-kordse lahjenduse ja lahjendamata DNA amplifitseerumise võimekust. CfDNA 10-kordse lahjenduse puhul saadi ΔCq väärtuseks 5,79 ning lahjendamata cfDNA puhul 2,34. FFPE DNA ΔCq väärtuseks 10-kordse lahjenduse puhul oli 10,16 ning lahjendamata FFPE DNA puhul 6,87. Saadud tulemuste põhjal otsustati 100-kordset lahjendust koematerjalist eraldatud genoomsele DNA-le mitte teha. Kuna cfDNA ja FFPE DNA proovide 10-kordseid lahjendusi võrreldi QCT 100-kordse lahjendusega, siis ei andnud see katse kinnitust, et antud DNA-d sobivad Illumina® TruSight® Tumor 26 protokolliga jaoks.

Sellele vaatamata prooviti koostada raamatukogu 4 proovist: tervete isikute plasmast eraldatud cfDNA segu, FFPE koelõigust eraldatud gDNA, 1% EGFR geeni c.2573T>G punktmutatsiooni ja 1% EGFR geeni c.2235_2249del15 deletsiooni kloonide *spike-in*-id tervete kontrollide cfDNA segusse. Raamatukogu olemasolu ja kvaliteeti kontrolliti TapeStation 2200 instrumendil. Õige pikkusega raamatukogu on vahemikus 300...330 bp, mida ei saavutatud ühegi katsetatud proovi puhul. Seejärel sooviti kontrollida, kas suurema DNA kontsentratsiooni korral võiks raamatukogu koostamine õnnestuda. Selleks kontsentreeriti *spike-in* segu CHRIST RVC 2-25 (Martin Christ Gefriertrocknungsanlagen GmbH, Saksamaa) vaakumkontsentratori abil 8 korda ning eraldati ka uus kogus tervete kontrollide rakuvaba DNA-d. Kõrgema kontsentratsiooni saamiseks elueeriti seekord DNA-d kaks korda. Saadud kontsentraadi ja cfDNA 100-kordse lahjenduse amplifitseerumise võimekust võrreldi samuti QCT kontrolliga ning saadi ΔCq väärtusteks vastavalt 7,04 ja 6,90. Kuigi need näidud olid suuremad kui tootja poolt soovitatud 4, otsustati nimetatud proovidest koostada raamatukogu, mille olemasolu kontrolliti TapeStation 2200 instrumendil. Õige pikkusega raamatukogu ei saadud ka sel korral. Seetõttu loobuti antud töös Illumina® poolt pakutava TruSight® Tumor 26 paneeli kasutamisest.

¹¹http://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_trusight_tumor.pdf

Amplikonipõhine teise põlvkonna sekveneerimine

Uuritavate geenide valik ja praimerite disain

Mitteväikerakulise kopsu kartsinoomiga isikute cfDNA ja FFPE DNA amplikonipõhiseks teise põlvkonna sekveneerimiseks valiti 5 ravimseoselist geeni: BRAF, EGFR, ERBB2, KRAS ja PIK3CA. Iga geeni puhul valiti 2 kõige sagedamini muteerunud eksonit, välja arvatud EGFR geen, mille puhul võeti analüüsimiseks türosiinkinaasi 4 eksonit. Praimereid disainides selgus, et EGFR geeni 20. eksonis leiduvatele mutatsioonidele pole võimalik etteantud parameetritega (materjal ja meetodika osa lehekülg 26) praimereid disainida. Otsustati proovida kahte varianti: üks praimeripaar amplifitseeriks pikemat kui 100 bp produkti ning kaks praimeripaari, mis amplifitseeriks lühemaid produkte. Kokku disainiti seega 14 praimeripaari. Materjal ja meetodika peatükis leheküljel 27 olevas tabelis 1 on toodud ainult kasutusse võetud praimerite lookusspetsiifilised järjestused. Disainitud praimerid jagati MultiPLX 2.1 (Kaplinski *et al.*, 2005) programmiga kaheks, kolmeks ja neljaks grupiks (tabel 4), mida võiks koos PCR-i reaktsioonis kasutada.

Tabel 4. MultiPLX 2.1 programmiga jagatud praimerigrupid. A. – praimerid jagatud nelja gruppi, B. – praimerid jagatud kolme gruppi, C. – praimerid jagatud kahte gruppi.

A.

Grupp 1	EGFRex18	EGFRex20	BRAFex15	BRAFex11
Grupp 2	KRASex2	ERBB2ex19	ERBB2ex20	PIK3CAex20
Grupp 3	EGFRex20mut2	EGFRex20mut1	KRASex3	
Grupp 4	PIK3CAex9	EGFRex19	EGFRex21	

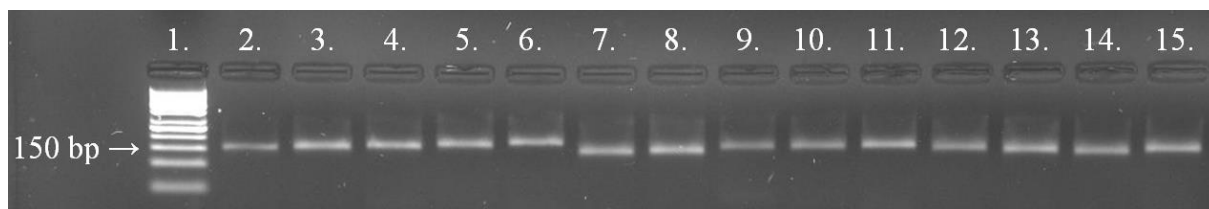
B.

Grupp 1	EGFRex18	BRAFex15	EGFRex20	EGFRex20mut2	BRAFex11
Grupp 2	EGFRex19	PIK3CAex9	ERBB2ex19	PIK3CAex20	EGFRex21
Grupp 3	EGFRex20mut1	KRASex3	ERBB2ex20	KRASex2	

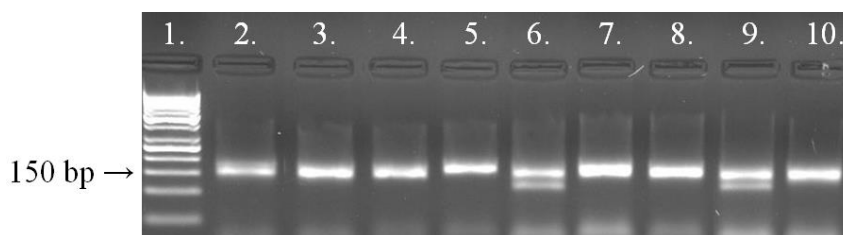
C.

Grupp 1	EGFRex20mut2	EGFRex18	EGFRex19	EGFRex20	BRAFex11	BRAFex15	PIK3CAex20
Grupp 2	PIK3CAex9	EGFRex20mut1	ERBB2ex19	ERBB2ex20	KRASex2	KRASex3	EGFRex21

Kõigepealt tuli veenduda üksikute ja seejärel grupeeritud praimeripaaride efektiivsuses. Selleks amplifitseeriti nimetatud regioone PCR-ga gDNA ja cfDNA pealt. PCR-i produktide olemasolu ning õiget pikkust kontrolliti 3% agarosgeelil (näitena gDNA geelipildid joonistel 13 ja 14). Produktide pikkus pidi jääma vahemikku 146...182 bp, mida oli ka näha visuaalsel kontrollil.



Joonis 13. Genoomse DNA pealt üksikute praimeripaaridega amplifitseeritud produktide visuaalne kontroll 3% agarosgeelil. 1. - O'GeneRuler 50 bp DNA suurusmarker (Thermo Fisher Scientific Inc.); 2.-15. - PCR-i produktid vasakult paremale: BRAF ekson 11 ja 15, EGFR ekson 18, 19, 20 (1), 20 (2), 20 ja 21, ERBB2 ekson 19 ja 20, KRAS ekson 2 ja 3, PIK3CA ekson 10 ja 21.



Joonis 14. Genoomse DNA pealt grupeeritud praimeripaaridega amplifitseeritud produktide visuaalne kontroll 3% agarosgeelil. 1. - O'GeneRuler 50 bp DNA suurusmarker; 2.-10. - PCR-i produktid vasakult paremale: jaotus A grupp 1, 2, 3 ja 4, jaotus B grupp 1, 2 ja 3, jaotus C grupp 1 ja 2 (vaata jaotusi tabelist 4.).

Visuaalne kontroll geelil näitas, et kõik praimeripaarid ja grupid töötavad enam-vähem võrdse efektiivsusega. Nähes, et PCR-i käigus paljundatakse ka pikemat EGFR geeni 20. eksoni produkti, otsustati loobuda kahest lühemat produkti amplifitseerivast praimeripaarist ning praimerid grupeeriti ümber (tabel 5).

Tabel 5. Ümber grupeeritud praimerigrupid. A. ja B. – praimerid jagatud kolme gruppi, C. – praimerid jagatud kahte gruppi.

A.

Grupp 1	EGFRex18	EGFRex20	BRAFex15	BRAFex11
Grupp 2	KRASex2	ERBB2ex19	ERBB2ex20	PIK3CAex20
Grupp 3	EGFRex19	EGFRex21	KRASex3	PIK3CAex9

B.

Grupp 1	EGFRex18	BRAFex15	EGFRex20	BRAFex11	
Grupp 2	EGFRex19	PIK3CAex9	ERBB2ex19	PIK3CAex20	EGFRex21
Grupp 3	KRASex3	ERBB2ex20	KRASex2		

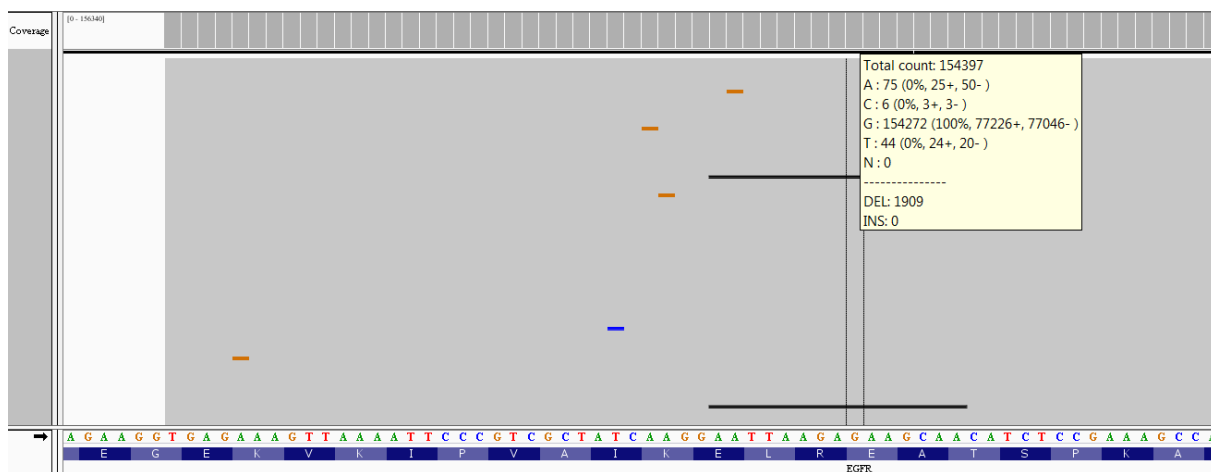
C.

Grupp 1	EGFRex18	EGFRex19	EGFRex20	BRAFex11	BRAFex15	PIK3CAex20
Grupp 2	PIK3CAex9	ERBB2ex19	ERBB2ex20	KRASex2	KRASex3	EGFRex21

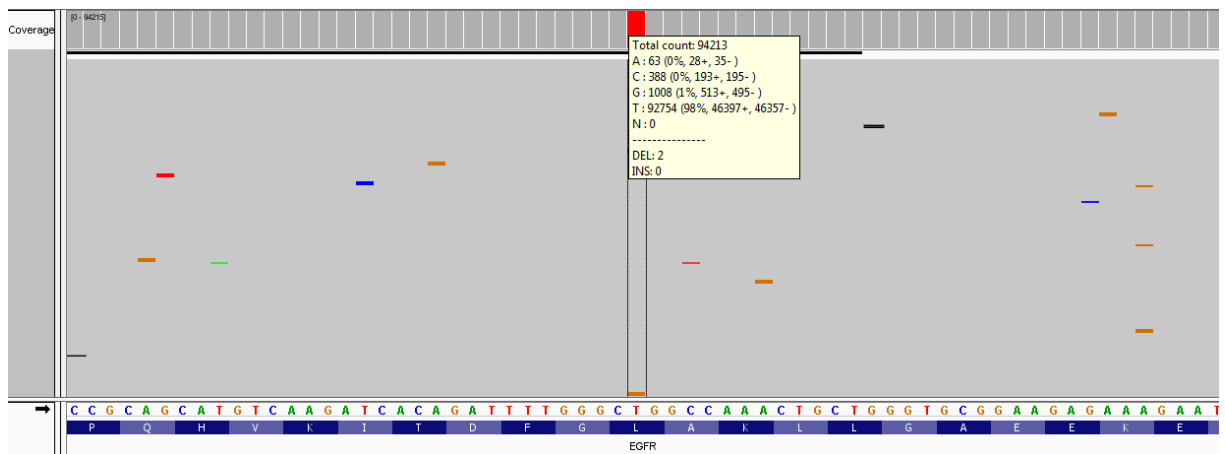
Et saada kindlust kõikide eksonite võrdses amplifitseerumises gruppide sees, tehti nendest raamatukogud teise põlvkonna sekveneerimise jaoks. Kuna jaotus A grupp 1 ning jaotus B grupp 1 sisaldavad samu praimeripaare, siis nende kahe grupi puhul tehti üks raamatukogu.

Sekveneerimistulemuste analüüs

Sekveneerimise tulemusena saadi kokku 19,6 miljonit 2 x 116 bp pikkust *paired-end* järjestust. 96,1% sekveneeritud järjestustest olid kvaliteediskooriga Q30 või rohkem. Proovi kohta saadi 0,26...0,47 miljonit järjestust, mis joondati GRCh37 inimese referentsgenoomile BWA-MEM algoritmi abil. Saadud .bam faile visualiseeriti IGV tarkvaraga ja kirjutati välja iga sekveneeritud ampliconi keskmine katvus, mille alusel võrreldi omavahel erinevate multipleks gruppide efektiivsust. CfDNA algmaterjali korral tehti 45 PCR-i tsükliga raamatukogu kaks korda, et võrrelda replikaatide sarnasust ja tõepoolest replikaatide ampliconide katvustes ei olnud tähelepanuväärseid erinevusi. Samuti polnud katvuste erinevused märgatavad 35 ja 45 tsükli jooksul paljundatud ampliconide vahel ning cfDNA ja FFPE DNA vahel. Mistõttu otsustati edaspidi kasutada multipleks amplicon PCR-i läbiviimiseks 35 tsüklit ning kuue ampliconiga multipleks gruppe (tabel 5 jaotus C). Nimetatud tarkvara kasutades määrati meetodi detektsioonipiiriks vähemalt 0,2%, kuna kõikide *spike-in* protsentide (0,2%, 0,5%, 1%, 2%, 4%, 8%) puhul oli näha nii deletsiooni kui ka punktmutatsiooni sisaldavaid järjestusi. Joonisel 15 on toodud deletsiooni 0,2% *spike-in* segu ja joonisel 16 punktmutatsiooni 0,2% *spike-in* segu sekveneeritud järjestused IGV tarkvaraga visualiseerimisel.



Joonis 15. Deletsiooni 0,2% *spike-in* segu NGS tulemuste analüüs IGV-s visualiseerituna. Alumisel real on näha EGFR geeni 19. eksoni aminohappeline järjestus, järgmisel real nukleotiidne järjestus. Ülemised vertikaalsed tulbad (*coverage*) tähistavad konkreetse nukleotiidse asukoha katvust. Mustad horisontaalsed jooned tähistavad deletsiooni esinemist. Kollasest kastist nähtub, et antud piirkonnas on järjestusi 154397 ning 1909 (1%) neist sisaldavad deletsiooni.



Joonis 16. Punktmutatsiooni 0,2% *spike-in* segu NGS tulemuste analüüs IGV-s visualiseerituna. Alumisel real on EGFR geeni 21. eksoni aminohappeline järjestus, järgmisel real nukleotiidne järjestus. Ülemised vertikaalsed tulbad (*coverage*) tähistavad konkreetse nukleotiidse asukoha katvust. Punasega on märgitud c.2573T>G mutatsiooni asukoht. Kollasest kastist nähtub, et selles asukohas on järjestusi 94213 ning 1008 (1%) neist sisaldavad T nukleotiidi asemel G nukleotiidi.

Uuritavad

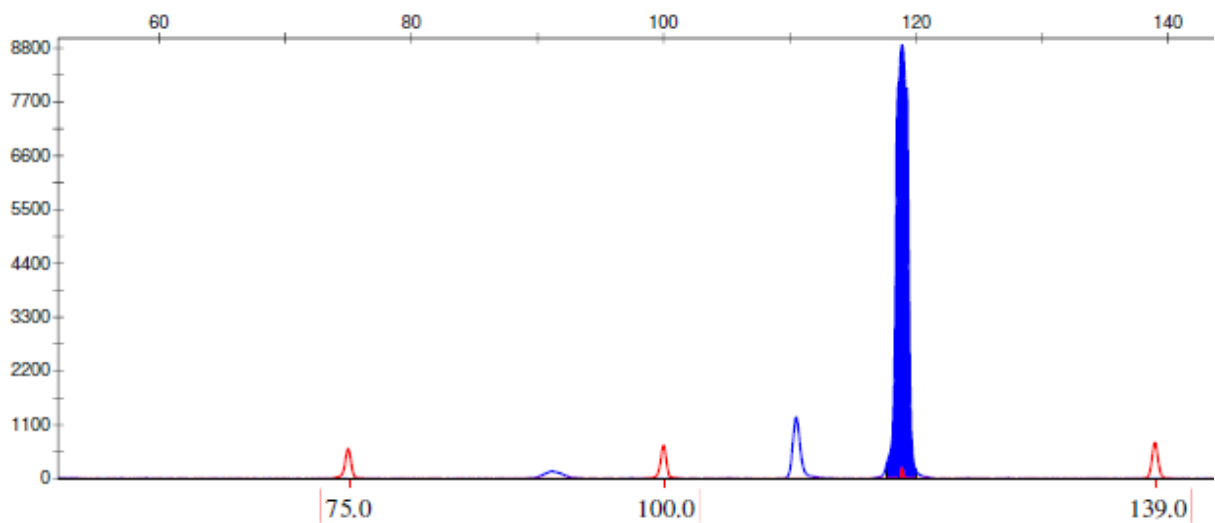
Uuritavate cfDNA ja FFPE DNA eraldamine

Uuritavate plasmast eraldatud esmase materjali rakuvaba DNA keskmine kontsentratsioon oli 1,28 ng/μl (vahemikus 0,14...31,40 ng/μl) ning progressioonil kogutud korduva materjali keskmine oli 0,85 ng/μl (vahemikus 0,21...1,41 ng/μl). Ühe uuritava (kA23) cfDNA eraldamisel saadi eluaadis väga kõrge DNA kontsentratsioon (31,40 ng/μl). See võis olla tingitud kas vererakkude lüüsumisest, hilisstaadiumi levinud kasvajast või juba alustatud ravist. FFPE koelõigust eraldatud genoomse DNA keskmine kontsentratsioon oli 91,29 ng/μl (vahemikus 0,28...484 ng/μl).

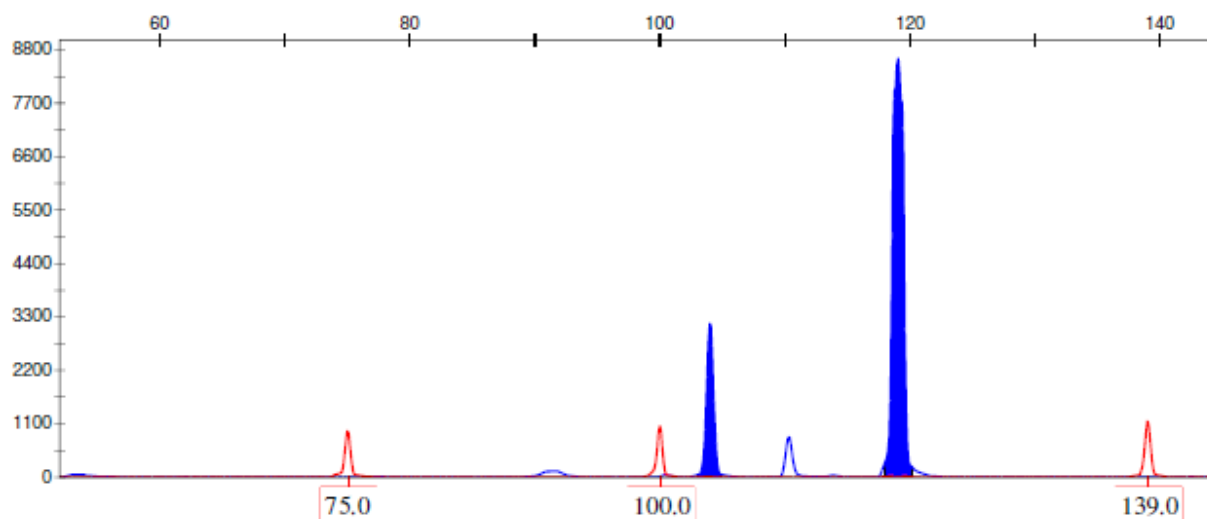
Ravimseoseliste mutatsioonide leiud

Fragmentanalüüs

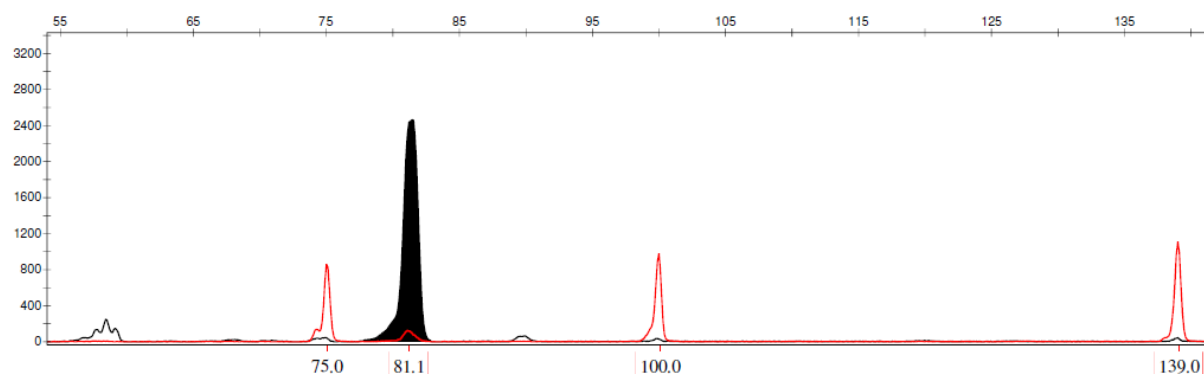
Fragmentanalüüsi tulemusena saadud elektroferogrammide põhjal otsustati, kas uuritaval esineb või ei esine otsitavat mutatsiooni. Deletsioon EGFR geeni 19. eksonis esines juhul, kui lisaks metsiktüüpi 121 bp fragmendile olin näha ka lühema pikkusega fragmenti. Punktmutatsiooni suhtes positiivseks loeti juhud, kui näha oli 67 bp pikkust fragmenti. Rakuvabas DNA-s esines EGFR geeni 19. eksonis deletsiooni 10 isikul (14%), kuid mitte ühelgi uuritaval ei esinenud 21. eksonis otsitavat punktmutatsiooni. Korduvast rakuvaba DNA materjalist ei leitud kumbagi mutatsiooni ühegi isiku puhul. FFPE koelõigust eraldatud DNA-s esines deletsioon 6 isikul (18%) ning punktmutatsioon 3 isikul (9%). Joonistel 17, 18, 19 ja 20 on näidetena toodud deletsioonide ning c.2573T>G punktmutatsiooni suhtes positiivsed ja negatiivsed elektroferogrammid.



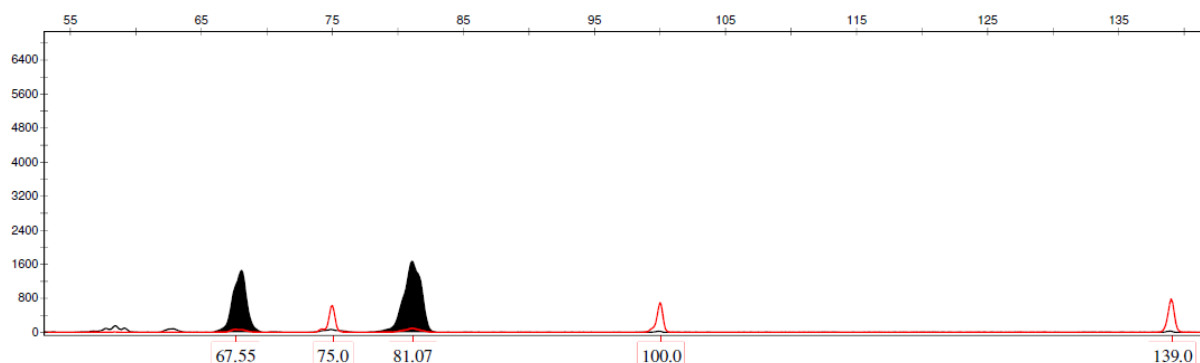
Joonis 17. Uuritava kA08 cfDNA elektroferogramm EGFR geeni 19. eksoni deletsioonide suhtes. X-teljel fragmentide pikkus aluspaarides (bp), y-teljel suhteline fluorestsents ühik (RFU). Punased piigid – pikkusmarkerid; sinisega täidetud piik – metsiktüüpi fragment. Sellel uuritaval ei esine cfDNA-s EGFR geeni 19. eksonis deletsioone.



Joonis 18. Uuritava kA02 cfDNA elektroferogramm EGFR geeni 19. eksoni deletsioonide suhtes. X-teljel fragmentide pikkus aluspaarides (bp), y-teljel suhteline fluorestsents ühik (RFU). Punased piigid – pikkusmarkerid; sinisega täidetud madalam piik – deletsiooniga fragment, sinisega täidetud kõrgem piik – metsiktüüpi fragment. Sellel uuritaval esineb cfDNA-s EGFR geeni 19. eksonis deletsioon.



Joonis 19. Uuritava kaA59 cfDNA elektroferogramm EGFR geeni 21. eksoni punktmutatsiooni suhtes. X-teljel fragmentide pikkus aluspaarides (bp), y-teljel suhteline fluorestsents ühik (RFU). Punased piigid – pikkusmarkerid; mustaga täidetud piik – metsiktüüpi fragment. Sellel uuritaval ei esine cfDNA-s c.2573T>G punktmutatsiooni.



Joonis 20. Uuritava kaA47 FFPE DNA elektroferogramm EGFR geeni 21. eksoni punktmutatsiooni suhtes. X-teljel fragmentide pikkus aluspaarides (bp), y-teljel suhteline fluorestsents ühik (RFU). Punased piigid – pikkusmarkerid; mustaga täidetud madalam piik – punktmutatsiooniga fragment; mustaga täidetud kõrgem piik – metsiktüüpi fragment. Sellel uuritaval esineb FFPE DNA-s c.2573T>G punktmutatsiooni.

Sekveneerimistulemused

Esimesel juhul saadi 19,66 miljonit 2 x 116 bp pikkust *paired-end* järjestust, millest 96% olid kvaliteediskooriga Q30 või rohkem ja teisel juhul saadi 16,14 miljonit 2 x 116 bp pikkust *paired-end* järjestust, millest 96,9% olid kvaliteediskooriga Q30 või rohkem. Esimese raamatukogu puhul saadi proovi kohta 0,04...0,30 ja teise raamatukogu puhul 0,04...0,33 miljonit järjestust.

Tabelis 6 on toodud IGV tarkvara kasutades tuvastatud EGFR geenis olevad 19. eksoni deletsioonide ja 21. eksoni c.2573T>G punktmutatsiooni leiud uuritavatel. 86% (19/22) leidudest olid 19. eksoni deletsioonid. 5 uuritava puhul oli nii cfDNA kui ka FFPE DNA materjalis sama mutatsioonileid. Uuritava kaA24 cfDNA-st leiti 19. eksoni deletsioon, kuid FFPE DNA-st 21. eksoni c.2573T>G punktmutatsioon. Ühel retsidiveerunud kasvaja uuringusse kaasatud isikul (kaA33) leiti lisaks ekson 19. deletsioonile cfDNA-st ka TKI sekundaarset resistentsust põhjustav p.T790M (c.2369C>T) mutatsioon 12% alleelse

fraktsiooniga. Uuritava koematerjalist nimetatud mutatsiooni ei tuvastatud. Selle põhjuseks on tõenäoliselt asjaolu, et koematerjal oli võetud 2 aastat varem kui vereproov ning vereproovi andmise eelselt oli uuritav saanud EGFR TKI ravi (erlotinib).

Tabel 6. IGV tarkvaraga tuvastatud uuritavate EGFR geeni mutatsioonileiud.

Materjal	Mutatsioon
cfDNA kA02	deletsioon
FFPE DNA kA02	deletsioon
cfDNA kA06	deletsioon
cfDNA kA09	deletsioon
FFPE DNA kA19	punktmutatsioon
cfDNA kA20	deletsioon
cfDNA kA24	deletsioon
FFPE DNA kA24	punktmutatsioon
cfDNA kA26	deletsioon
FFPE DNA kA27	deletsioon
cfDNA kA28	deletsioon
FFPE DNA kA28	deletsioon
cfDNA kA33	deletsioon
FFPE DNA kA33	deletsioon
FFPE DNA kA47	punktmutatsioon
cfDNA kA48	deletsioon
FFPE DNA kA48	deletsioon
FFPE DNA kA54(2)	deletsioon
cfDNA kA59	deletsioon
FFPE DNA kA59	deletsioon
cfDNA kA62	deletsioon
cfDNA kA63	deletsioon

Indelite tuvastamiseks kasutatud aiHunter tarkvara abil saadud tulemused on toodud tabelis 7. 13 uuritava puhul leiti analüüsitud materjalidest EGFR geeni 19. eksoni deletsioone, kuid

mitte ühelgi uuritaval polnud insertioone. 38% (5/13) deletsioonidest esines nii cfDNA-s kui ka FFPE DNA-s. Lisaks nendele leiti kolmest proovist (cfDNA kA02, cfDNA kA24 ja FFPE DNA kA54 (2)) ka teine, pikem 30...54 nukleotiidiline deletsioon, mida pole tabelis välja toodud.

Tabel 7. AiHunter tarkvaraga tuvastatud indelid. Chr – kromosoom. Positsioon – indeli koordinaadid kromosoomil. Alt – alternatiivne järjestus. Katvus – amplikoni kogu katvus. Sagedus – indelis sisalduva järjestuse esinemissagedus.

Materjal	Chr	Positsioon	Referents järjestus	Alt	Katvus	Sagedus
cfDNA kA02	7	55242440	AGGAATTAAGAGAAGC	A	21956	8%
cfDNA kA06	7	55242442	GAATTAAGAGAAGCAA	G	40412	4%
cfDNA kA09	7	55242442	GAATTAAGAGAAGCAA	G	37650	1%
cfDNA kA20	7	55242442	GAATTAAGAGAAGCAACAT	G	32748	1%
cfDNA kA24	7	55242442	GAATTAAGAGAAGCAA	G	10932	1%
cfDNA kA26	7	55242445	TTAAGAGAAGCAACATCTC	T	41954	29%
cfDNA kA28	7	55242441	GGAATTAAGA	G	70364	48%
cfDNA kA33	7	55242443	AATTAAGAGAAGCAAC	A	70678	28%
cfDNA kA48	7	55242440	AGGAATTAAGAGAAGC	A	37318	9%
cfDNA kA59	7	55242445	TTAAGAGAAGCAACATCTC	T	34144	2%
cfDNA kA62	7	55242441	GGAATTAAGA	G	36862	9%
cfDNA kA63	7	55242441	GGAATTAAGAGAAGCA	G	34034	9%
FFPE DNA kA02	7	55242440	AGGAATTAAGAGAAGC	A	62414	46%
FFPE DNA kA27	7	55242440	AGGAATTAAG	A	70244	10%
FFPE DNA kA28	7	55242441	GGAATTAAGA	G	83324	77%
FFPE DNA kA33	7	55242443	AATTAAGAGAAGCAAC	A	89374	90%
FFPE DNA kA48	7	55242440	AGGAATTAAGAGAAGC	A	48306	63%
FFPE DNA kA54(2)	7	55242442	GAATTAAGAGAAGCAA	G	70942	1%
FFPE DNA kA59	7	55242445	TTAAGAGAAGCAACATCTC	T	41472	16%

GATK programmiga tuvastatud anoteeritud variandid on toodud tabelis 8. Ühenukleotiidilisi asendusi leiti EGFR ja KRAS geenidest. Ühel uuritaval esines sama mutatsioon nii cfDNA kui ka FFPE DNA materjalis. Ülejäänud 9 isiku puhul leidis nukleotiidivahetusi kas cfDNA või FFPE DNA materjalis. Fragmentanalüüsil otsitavat EGFR geeni 21. eksoni c.2573T>G punktmutatsiooni esines kahe uuritava (kA19 ja kA47) FFPE materjalis.

Tabel 8. GATK programmiga tuvastatud ühenukleotiidsed muutused. Ref – referents järjestus. Alt – alternatiivne järjestus. AH – aminohape. ID - *reference* SNP ID (rs) number.

Materjal	Geen	Ekson	Ref	Alt	Valgu positsioon	AH	ID
cfDNA kA26	EGFR	19	C	T	755	A	rs727503017
cfDNA kA28	EGFR	19	G	C	750	A/P	rs121913229
cfDNA kA62	EGFR	19	G	C	750	A/P	rs121913229
FFPE DNA kA16	EGFR	21	T	A	861	L/Q	rs121913444
FFPE DNA kA19	EGFR	21	T	G	858	L/R	rs121434568
FFPE DNA kA25	KRAS	2	C	A	12	G/V	rs121913529
FFPE DNA kA27	EGFR	19	G	C	750	A/P	rs121913229
FFPE DNA kA27	EGFR	19	A	C	750	A	COSM1451590
FFPE DNA kA28	EGFR	19	G	C	750	A/P	rs121913229
FFPE DNA kA47	EGFR	21	T	G	858	L/R	rs121434568
FFPE DNA kA51	KRAS	2	C	A	12	G/C	rs121913530
FFPE DNA kA55	KRAS	2	C	G	12	G/A	rs121913529

Mutatsioonileidude võrdlemine

Diagnostikalabori kui standardi põhjal nelja klassi (õige positiivne ja negatiivne ning valepositiivne ja –negatiivne) jaotatud 28 uuritava EGFR geeni mutatsioonileidude tulemused on toodud tabelis 9 ning kolme analüüsimeetodi tundlikkus, spetsiifilisus, positiivne ennustusväärtus (PPV) ja kokkulangevus diagnostikalabori tulemustega on toodud tabelis 10.

Tabel 9. EGFR geeni 19. eksoni deletsioonide või 21. eksoni c.2573T>G ühenukleotiidilise asenduse mutatsioonileidude tõesuse hinnang diagnostikalabori tulemuste põhjal. TP – õige positiivne, TN – õige negatiivne, FP – valepositiivne, FN – valenegatiivne.

Uuritav	Diagn lab	Fragmentanalüüs		NGS, manuaalne		NGS, tarkvara	
	FFPE DNA	FFPE DNA	cfDNA	FFPE DNA	cfDNA	FFPE DNA	cfDNA
kA02	+	TP	TP	TP	TP	TP	TP
kA05	-	TN	TN	TN	TN	TN	TN
kA08	-	TN	TN	TN	TN	TN	TN
kA12	-	TN	TN	TN	TN	TN	TN
kA16	-	TN	TN	TN	TN	TN	TN
kA19	+	TP	FN	TP	FN	TP	FN
kA24	+	TP	FN	TP	FN	FN	FN
kA25	-	TN	TN	TN	TN	TN	TN
kA27	-	FP	TN	FP	TN	FP	TN
kA28	+	TP	TP	TP	TP	TP	TP
kA32	-	TN	TN	TN	TN	TN	TN
kA33	+	TP	TP	TP	TP	TP	TP
kA40	-	TN	TN	TN	TN	TN	TN
kA41	-	TN	TN	TN	TN	TN	TN

kA44	-	TN	TN	TN	TN	TN	TN
kA45	-	TN	TN	TN	TN	TN	TN
kA46	-	TN	TN	TN	TN	TN	TN
kA47	+	TP	FN	TP	FN	TP	FN
kA48	+	TP	TP	TP	TP	TP	TP
kA49	-	TN	TN	TN	TN	TN	TN
kA50	-	TN	TN	TN	TN	TN	TN
kA51	-	TN	TN	TN	TN	TN	TN
kA52	-	TN	TN	TN	TN	TN	TN
kA53	-	TN	TN	TN	TN	TN	TN
kA54	-	TN	TN	FP	TN	FP	TN
kA55	-	TN	TN	TN	TN	TN	TN
kA58	-	TN	TN	TN	TN	TN	TN
kA59	+	TP	TP	TP	TP	TP	TP

Tabel 10. Kolme analüüsimeetodi tundlikkus, spetsiifilisus, positiivne ennustusväärtus (PPV) ja kokkulangevus diagnostikalabori tulemustega.

	Fragmentanalüüs		NGS, manuaalne		NGS, tarkvara	
	FFPE DNA	cfDNA	FFPE DNA	cfDNA	FFPE DNA	cfDNA
Tundlikkus	100%	63%	100%	63%	88%	63%
Spetsiifilisus	95%	100%	90%	100%	90%	100%
PPV	89%	100%	80%	100%	78%	100%
Kokkulangevus	96%	89%	93%	89%	89%	89%

Diagnostikalaboris testitud 28-st FFPE koematerjalist esines EGFR geeni mutatsioon 8 uuritavaal (29%). FFPE DNA puhul andis fragmentanalüüsi meetodil mutatsioonide tuvastamine võrreldes NGS meetoditega enam kokkulangevaid (96%) tulemusi diagnostikalaboriga. Eeldusel, et diagnostikalaboris tuvastatud leiud olid 100% tõesed ja ei esinenud ühtki valepositiivset ega –negatiivset tulemust, andis fragmentanalüüsi ja NGS võrdluses kõikidest positiivsetest leidudest enam õigeid positiivseid tulemusi (PPV) samuti fragmentanalüüs (89%).

Tabelis 11 on toodud kõigi uuritavate rakuvaba DNA testimisel fragmentanalüüsi ja teise põlvkonna sekveneerimise meetoditega EGFR geeni kahe peamise mutatsiooni tuvastamise tundlikkus, spetsiifilisus ja positiivne ennustusväärtus.

Tabel 11. Kõigi 70 uuritava cfDNA EGFR geeni mutatsioonide testimise tundlikkus, spetsiifilisus ja positiivne ennustusväärtus (PPV).

	Fragmentanalüüs	NGS, manuaalne ja tarkvarapõhine
	cfDNA	cfDNA
Tundlikkus	64%	79%
Spetsiifilisus	98%	100%
PPV	90%	100%

Kõigi 70 uuritava cfDNA testimise tundlikkus ja spetsiifilisus oli NGS meetodil kõrgem kui fragmentanalüüsi meetodil. Kõik NGS meetodil tuvastatud leidud osutusid õigeteks positiivseteks (PPV = 100%).

Iga meetodi puhul võrreldi omavahel ka rakuvabast ning koematerjalist tuvastatud EGFR geeni mutatsioonileidude kokkulangevust (tabelid 12, 13 ja 14).

Tabel 12. Fragmentanalüüsi meetodil tuvastatud EGFR geeni mutatsioonileidude kokkulangevus cfDNA ja FFPE DNA materjalis.

cfDNA	FFPE DNA		Kokku
	+	-	
+	5	0	5
-	4	24	28
Kokku	9	24	33 (100%)

Tabel 13. NGS manuaalse analüüsimetodil tuvastatud EGFR geeni mutatsioonileidude kokkulangevus cfDNA ja FFPE DNA materjalis.

cfDNA	FFPE DNA		Kokku
	+	-	
+	5	0	5
-	5	23	28
Kokku	10	23	33 (100%)

Tabel 14. NGS tarkvarapõhise analüüsimetodil tuvastatud EGFR geeni mutatsioonileidude kokkulangevus cfDNA ja FFPE DNA materjalis.

cfDNA	FFPE DNA		Kokku
	+	-	
+	5	0	5
-	4	24	28
Kokku	9	24	33 (100%)

Fragmentanalüüsi ja NGS tarkvarapõhise analüüsimetodi korral oli 33 uuritava võrreldavast materjalist 29 puhul (88%) cfDNA ja FFPE DNA tulemused kokkulangevad nii

mutatsioonileiu kui ka selle puudumise osas. NGS manuaalse analüüsimeetodi korral ühtisid mutatsioonileid ja selle puudumine 28 (85%) materjali puhul.

Arutelu

Vähkkasvajate molekulaarne iseloomustamine võib olla oluline nii varasel haiguse avastamisel, ravi valikul, residuaalhaiguse jälgimisel ja resistentsuse tõttu progressiooni tekkimisel ravimeetodi vahetamisel. Uued sihtmärkravimid on parandanud ravivastust ja võrreldes kemoteraapiaga vähendanud kõrvaltoimete esinemist, kuid annavad positiivse ravivastuse ainult kindlate mutatsioonide olemasolul ja resistentsusmutatsioonide puudumisel. Molekulaargeneetilise analüüsi teeb aga keerukaks tuumori rakkudevaheline heterogeensus ning hetkel peamiselt biopsiamaterjali põhjal raviplaani koostamine, mis annab infot ainult konkreetset ajahetket konkreetsetes kohas esinevatest mutatsioonidest. Käesoleva magistritöö eesmärgiks oligi välja töötada molekulaardiagnostika meetodid, millega oleks võimalik tuvastada ravimseoseliste mutatsioonide esinemist kasvajatega isikute tsirkuleerivast rakuvabast DNA-st ning võrrelda tulemusi cfDNA ja fikseeritud koematerjali kasutamisel.

Esmalt tuli leida parim viis vereplasmas tsirkuleeriva rakuvaba DNA eraldamiseks, mis annaks suurima saagikuse. Selleks kasutati kahte kommertsiaalset kitti, millega eraldatud DNA kontsentratsioone võrreldi omavahel. QIAamp[®] CNA kitiga eraldatud uuritavate rakuvaba DNA keskmine saagikus ja kontsentratsioonid on võrreldavad varasemalt kirjanduses avaldatuga ja seega otsustati nimetatud kiti kasuks.

Mutatsiooniuuringutes kasutati amplifitseeritud DNA fragmentanalüüsi ning NGS analüüsi. FFPE koelõikudest eraldatud DNA kasvajaspetsiifiliste geenide sekveneerimiseks mõeldud Illumina[®] TruSight[®] Tumor 26 paneeli abil raamatukogu loomine ei tulnud tõenäoliselt välja seetõttu, et DNA algkogus polnud piisav 174 amplikoni paljundamiseks. Alternatiivse lahendusena töötati välja amplikonipõhine NGS metoodika EGFR signaaliraja 5 geeni sagedasemate mutatsioonide määramiseks nii cfDNA-st kui ka FFPE DNA-st.

Uuritavate piirkondade DNA amplifikatsioon toimus edukalt ka multipleks reaktisoonis. Kuigi sekveneerimisel saavutati amplikonide kõrge katvus, tuleb täiendavat tööd teha variantide automaatse leidmise algoritmidega, kuna teadaolevalt on laialt levinud lahendustel raskusi variantide tuvastamisega madala muteerunud DNA osakaalu puhul. Võrreldes punktmutatsioonidega on probleemsem indelide tuvastamine. Algoritmide optimeerimiseks kasutati paralleelselt NGS tulemuste manuaalset kontrolli IGV (Robinson *et al.*, 2011)

tarkvaraga. Selle töömahukuse tõttu on antud magistritöö NGS tulemustes kajastatud vaid EGFR geeni mutatsioonianalüüs ning selle võrdlus DNA fragmentanalüüsi tulemustega.

Esmalt hinnati tulemusi 28 võrreldava cfDNA ja FFPE materjali puhul, mis olid eelnevalt EGFR mutatsioonide osas uuritud kahes erinevas sertifitseeritud diagnostikalaboris. Diagnostiliste uuringute tulemused ei pärine otse diagnostikalaborist, vaid on kogutud koos kliinilise andmestikuga raviarstide käest.

Diagnostikalaboritega võrreldes lahknenuid tulemustest on võimalik, et väljatöötatud meetoditega esialgselt valepositiivseks hinnatud tulemused uuritavate kA27 ja kA54 FFPE DNA puhul on siiski õiged positiivsed, kuna mutantse DNA osakaal oli NGS analüüsimeetodiga mõõdetuna väga madal (umbes 1% kogu järjestustest) ning seetõttu võis diagnostikalaborites olla raskusi nende usaldusväärse tuvastamisega. Kuna nendel juhtudel ei alustatud ravi türosiinkinaasi inhibiitoritega, ei saa mutatsiooni esinemise kohta järeldust teha ka ravitundlikkuse põhjal. Valenegatiivsed tulemused uuritavate kA19, kA24 ja kA47 cfDNA materjalis võivad viidata muteerunud DNA väga madalale osakaalule proovis. Metaanalüüsides on teada, et cfDNA analüüsimisel on võrreldes koematerjaliga väiksem tundlikkus. Samuti ei pruugi tuumoritele omase heterogeensuse tõttu igast kasvajakarakust DNA-d vereringesse sattuda. Kuigi projekti käigus oli planeeritud koguda vereproove enne patsientide ravi alustamist, võis siiski juhtuda, et uuritav kA19 oli juba TKI raviga alustanud ja see võib oluliselt vähendada kasvajast pärineva cfDNA osakaalu. Uuritava kA24 FFPE koematerjalist leiti nii diagnostikalaboris testides kui ka fragmentanalüüsi ja NGS manuaalse analüüsi käigus 21. eksoni punktmutatsioon. Seevastu cfDNA materjalist leiti NGS tarkvarapõhise ja manuaalse analüüsi käigus umbes 1% 19. eksoni deletsiooniga järjestusi. Eeldusel, et ükski leid pole valepositiivne võib selline tulemus näidata hoopis kasvaja klonaalset heterogeensust.

Kõige selgemalt näitab cfDNA analüüsist saadavat kasu EGFR TKI sekundaarset resistentsust põhjustava p.T790M mutatsiooni leid uuritaval kA33. Antud isikul oleks näidustatud esimese põlvkonna TKI asendamine kolmanda põlvkonna ravimiga.

KOKKUVÕTE

Käesoleva magistritöö kirjanduse osas anti ülevaade üldisest onkogeneetikast, spetsiifilisemalt kopsuvähist ning seda põhjustavatest juhtmutatsioonidest ja nende ravitundlikkusest. Lisaks kirjeldati hetkel laialdaselt uuritavat vedela biopsia valdkonda ja selle kasutamisevõimalusi. Samuti kirjeldati lähemalt teise põlvkonna sekveneerimist, mida kasutati ka töö eksperimentaalses osas.

Töö praktilise osa eesmärgiks oli testida tsirkuleerivat rakuvaba DNA-d ja FFPE koematerjalist eraldatud DNA-d kopsuvähi puhul sagedamini muteerunud geenides ravitundlikkust mõjutavate mutatsioonide suhtes. Esmalt töötati välja EGFR geeni kahe enamlevinud mutatsiooni (19. eksoni deletsioonide ja 21. eksoni punktmutatsiooni) tuvastamiseks alleel-spetsiifilise fragmentanalüüsi protokoll ja kopsuvähiga seostatud 5 geeni (EGFR, BRAF, ERBB2, KRAS, PIK3CA) amplikonipõhise teise põlvkonna sekveneerimise töövoog. Väljatöötatud meetoditega on võimalik tuvastada mutantset alleeli, mille osakaal on alla 1%.

Neid meetodeid kasutades analüüsiti 70 kopsuvähiga patsiendi plasmast eraldatud rakuvaba DNA-d ja FFPE koelõigust eraldatud genoomset DNA-d. Leide analüüsides käsitleti esmalt kahe sertifitseeritud diagnostikalabor tulemusi absoluutse tõena. Hiljem, analüüsides kõiki koeproove, käsitleti tõesena vähemalt kahe meetodiga tuvastatud tulemusi. Selle põhjal arvutati fragmentanalüüsi, teise põlvkonna sekveneerimise manuaalse ja tarkvarapõhise analüüsi tundlikkus ning spetsiifilisus. Mutatsioonanalüüs FFPE DNA-st on kõrgema tundlikkusega kui cfDNA-st, kuid rakuvaba DNA eeliseks on kõrge spetsiifilisus mutatsioonide tuvastamisel.

Uuringusse kaasatud mitteväikerakulise kopsuvähiga patsientide rakuvaba DNA sekveneerimisel NGS meetodiga tuvastati ühel juhul ka esimese põlvkonna EGFR-TKI sekundaarset resistentsust põhjustav mutatsioon p. T790M, mida diagnostikalaborid polnud suutnud koematerjalist leida. Ka NGS meetodil ei tuvastatud esmasel diagnoosil võetud koematerjalist antud mutatsiooni. Leid esines EGFR-TKI ravi järgselt. Antud isikul oleks näidustatud esimese põlvkonna TKI ravi asendamine kolmanda põlvkonna TKI-ga.

Pikemas plaanis on kavas paneelanalüüs põhjalikumalt juurutada. Samuti kaalutakse paneelis geenide valikud laiendada. Kliinilisi andmeid kasutades soovime uurida seoseid mutatsioonide, ravitundlikkuse, progressioonivaba elulemuse ja üldelulemusega. Töö autor

usub, et cfDNA kasutamine muutub mitmete vähkkasvajate puhul peagi rutiiniks ka kliinilises praktikas.

Detecting cancer related mutations in cell-free DNA of lung cancer patients

Madli Tamm

Summary

According to WHO, lung cancer is the leading cause of cancer deaths worldwide. Besides invasive tumor biopsy, molecular diagnostic analysis of tumor cell-free DNA (cfDNA) is becoming increasingly popular as a method enabling to capture the whole tumor heterogeneity in almost real-time setting. This may have applications in both early diagnostics, as well as therapeutic monitoring of alterations predictive for drug response.

An allele-specific fragment length analysis (FLA) workflow for EGFR gene two common mutations (exon 19 deletions and exon 21 L858R mutation) and an amplicon-based multiplex next-generation sequencing (NGS) workflow for 5 lung cancer-related genes (EGFR, BRAF, ERBB2, KRAS, PIK3CA) was set up. Plasma samples from 70 lung carcinoma patients with 33 matching FFPE materials were analysed using these methods. Test results by two major certified diagnostic laboratories were considered to be the gold standard and compared to the results of FLA and NGS.

Results show that optimized workflows for FLA and NGS approaches can detect mutant allele content less than 1%. The overall concordance between diagnostic laboratory results and FLA, NGS manual analysis and NGS software based analysis results was 93%, 91% and 89%, respectively. The sensitivity of analysing FFPE DNA material was better than analysing cfDNA but cfDNA had better specificity over FFPE DNA material. In addition to the EGFR exon 19 deletions and ex 21 L858R pointmutations a T790M resistance mutation was found in one female patient. This could mean that her current treatment plan should be changed to third-generation EGFR TKI-s.

The further aim of this project is to perform a larger screen of drug susceptibility and resistance mutations in the study cohort, involving detection of copy number variations and DNA methylation changes. Author strongly believes that cfDNA analysis will become a clinical routine in various cancers in the near future.

TÄNUAVALDUSED

Esmalt sooviksin tänada oma juhendajat Neeme Tõnissoni, kes võimaldas mul teha magistritööd niivõrd põneval teemal. Suur aitäh ka minu kaasjuhendajatele Paula Annile ja Mardile. Samuti tänaksin kõiki „Tsirkuleeriv DNA kopsuvähi biomarkerina“ projektis osalevaid isikuid, eriti Katrinit.

KASUTATUD KIRJANDUS

- Alberg, A.J. & Samet, J.M. (2003) Epidemiology of lung cancer. *Chest*, **123**, 21S-49S.
- Atamaniuk, J., Vidotto, C., Tschan, H., Bachl, N., Stuhlmeier, K.M. & Müller, M.M. (2004) Increased concentrations of cell-free plasma DNA after exhaustive exercise. *Clin Chem*, **50**, 1668-1670.
- Ballaz, S. & Mulshine, J.L. (2003) The potential contributions of chronic inflammation to lung carcinogenesis. *Clin Lung Cancer*, **5**, 46-62.
- Bentley, D.R. & Balasubramanian, S. & Swerdlow, H.P. & Smith, G.P. & Milton, J. & Brown, C.G. & Hall, K.P. & Evers, D.J. & Barnes, C.L. & Bignell, H.R. & Boutell, J.M. & Bryant, J. & Carter, R.J. & Keira Cheetham, R. & Cox, A.J. & Ellis, D.J. & Flatbush, M.R. & Gormley, N.A. & Humphray, S.J. & Irving, L.J. & Karbelashvili, M.S. & Kirk, S.M. & Li, H. & Liu, X. & Maisinger, K.S. & Murray, L.J. & Obradovic, B. & Ost, T. & Parkinson, M.L. & Pratt, M.R. & Rasolonjatovo, I.M. & Reed, M.T. & Rigatti, R. & Rodighiero, C. & Ross, M.T. & Sabot, A. & Sankar, S.V. & Scally, A. & Schroth, G.P. & Smith, M.E. & Smith, V.P. & Spiridou, A. & Torrance, P.E. & Tzonev, S.S. & Vermaas, E.H. & Walter, K. & Wu, X. & Zhang, L. & Alam, M.D. & Anastasi, C. & Aniebo, I.C. & Bailey, D.M. & Bancarz, I.R. & Banerjee, S. & Barbour, S.G. & Baybayan, P.A. & Benoit, V.A. & Benson, K.F. & Bevis, C. & Black, P.J. & Boodhun, A. & Brennan, J.S. & Bridgham, J.A. & Brown, R.C. & Brown, A.A. & Buermann, D.H. & Bundu, A.A. & Burrows, J.C. & Carter, N.P. & Castillo, N. & Chiara E Catenazzi, M. & Chang, S. & Neil Cooley, R. & Crake, N.R. & Dada, O.O. & Diakoumakos, K.D. & Dominguez-Fernandez, B. & Earnshaw, D.J. & Egbujor, U.C. & Elmore, D.W. & Etchin, S.S. & Ewan, M.R. & Fedurco, M. & Fraser, L.J. & Fuentes Fajardo, K.V. & Scott Furey, W. & George, D. & Gietzen, K.J. & Goddard, C.P. & Golda, G.S. & Granieri, P.A. & Green, D.E. & Gustafson, D.L. & Hansen, N.F. & Harnish, K. & Haudenschild, C.D. & Heyer, N.I. & Hims, M.M. & Ho, J.T. & Horgan, A.M., et al. (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, **456**, 53-59.
- Berg, M. & Soreide, K. (2012) EGFR and downstream genetic alterations in KRAS/BRAF and PI3K/AKT pathways in colorectal cancer: implications for targeted therapy. *Discov Med*, **14**, 207-214.
- Bordi, P., Del Re, M., Danesi, R. & Tiseo, M. (2015) Circulating DNA in diagnosis and monitoring EGFR gene mutations in advanced non-small cell lung cancer. *Transl Lung Cancer Res*, **4**, 584-597.
- Brambilla, E. & Gazdar, A. (2009) Pathogenesis of lung cancer signalling pathways: roadmap for therapies. *Eur Respir J*, **33**, 1485-1497.
- Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B. & Cohen, S.M. (2003) bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. *Cell*, **113**, 25-36.
- Calin, G.A., Sevignani, C., Dumitru, C.D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M. & Croce, C.M. (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci U S A*, **101**, 2999-3004.
- Cantley, L.C. (2002) The phosphoinositide 3-kinase pathway. *Science*, **296**, 1655-1657.
- Chen, X., Wang, X., He, H., Liu, Z., Hu, J.F. & Li, W. (2015) Combination of circulating tumor cells with serum carcinoembryonic antigen enhances clinical prediction of non-small cell lung cancer. *PLoS One*, **10**, e0126276.
- Chen, Y., Shi, J.X., Pan, X.F., Feng, J. & Zhao, H. (2013) Identification of candidate genes for lung cancer somatic mutation test kits. *Genet Mol Biol*, **36**, 455-464.

- Cheng, Y.W., Chiou, H.L., Sheu, G.T., Hsieh, L.L., Chen, J.T., Chen, C.Y., Su, J.M. & Lee, H. (2001) The association of human papillomavirus 16/18 infection with lung cancer among nonsmoking Taiwanese women. *Cancer Res*, **61**, 2799-2803.
- Cooper, D.N. (2005) *The molecular genetics of lung cancer*, Springer, Berlin ; New York.
- Couraud, S., Vaca-Paniagua, F., Villar, S., Oliver, J., Schuster, T., Blanché, H., Girard, N., Trédaniel, J., Guilleminault, L., Gervais, R., Prim, N., Vincent, M., Margery, J., Larivé, S., Foucher, P., Duvert, B., Vallee, M., Le Calvez-Kelm, F., McKay, J., Missy, P., Morin, F., Zalcman, G., Olivier, M., Souquet, P.J. & investigators, B.I.-. (2014) Noninvasive diagnosis of actionable mutations by deep sequencing of circulating free DNA in lung cancer from never-smokers: a proof-of-concept study from BioCAST/IFCT-1002. *Clin Cancer Res*, **20**, 4613-4624.
- Devarakonda, S., Morgensztern, D. & Govindan, R. (2015) Genomic alterations in lung adenocarcinoma. *Lancet Oncol*, **16**, e342-351.
- DeVita, V.T., Lawrence, T.S. & Rosenberg, S.A. (2015a) *DeVita, Hellman, and Rosenberg's Cancer, Principles & Practice of Oncology*, Wolters Kluwer, Philadelphia.
- DeVita, V.T., Lawrence, T.S. & Rosenberg, S.A. (2015b) *DeVita, Hellman, and Rosenberg's Cancer, Principles & Practice of Oncology, 10th edition*, Wolters Kluwer, Philadelphia.
- Devonshire, A.S., Whale, A.S., Gutteridge, A., Jones, G., Cowen, S., Foy, C.A. & Huggett, J.F. (2014) Towards standardisation of cell-free DNA measurement in plasma: controls for extraction efficiency, fragment size bias and quantification. *Anal Bioanal Chem*, **406**, 6499-6512.
- Fearon, E.R. (1997) Human cancer syndromes: clues to the origin and nature of cancer. *Science*, **278**, 1043-1050.
- Forbes, S.A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H., Ding, M., Bamford, S., Cole, C., Ward, S., Kok, C.Y., Jia, M., De, T., Teague, J.W., Stratton, M.R., McDermott, U. & Campbell, P.J. (2015) COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res*, **43**, D805-811.
- Fraga, M.F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K., Iyer, N.G., Pérez-Rosado, A., Calvo, E., Lopez, J.A., Cano, A., Calasanz, M.J., Colomer, D., Piris, M.A., Ahn, N., Imhof, A., Caldas, C., Jenuwein, T. & Esteller, M. (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet*, **37**, 391-400.
- Francisci, S., Minicozzi, P., Pierannunzio, D., Ardanaz, E., Eberle, A., Grimsrud, T.K., Knijn, A., Pastorino, U., Salmerón, D., Trama, A., Sant, M. & Group, E.-W. (2015) Survival patterns in lung and pleural cancer in Europe 1999-2007: Results from the EUROCORE-5 study. *Eur J Cancer*.
- Frenel, J.S., Carreira, S., Goodall, J., Roda, D., Perez-Lopez, R., Tunariu, N., Riisnaes, R., Miranda, S., Figueiredo, I., Nava-Rodrigues, D., Smith, A., Leux, C., Garcia-Murillas, I., Ferraldeschi, R., Lorente, D., Mateo, J., Ong, M., Yap, T.A., Banerji, U., Gasi Tandefelt, D., Turner, N., Attard, G. & de Bono, J.S. (2015) Serial Next-Generation Sequencing of Circulating Cell-Free DNA Evaluating Tumor Clone Response To Molecularly Targeted Drug Administration. *Clin Cancer Res*, **21**, 4586-4596.
- Friedberg, E.C. (2003) DNA damage and repair. *Nature*, **421**, 436-440.
- Froyen, G., Broekmans, A., Hillen, F., Pat, K., Achten, R., Mebis, J., Rummens, J.L., Willemse, J. & Maes, B. (2016) Validation and Application of a Custom-Designed Targeted Next-Generation Sequencing Panel for the Diagnostic Mutational Profiling of Solid Tumors. *PLoS One*, **11**, e0154038.
- Garces, Y.I., Yang, P., Parkinson, J., Zhao, X., Wampfler, J.A., Ebbert, J.O. & Sloan, J.A. (2004) The relationship between cigarette smoking and quality of life after lung cancer diagnosis. *Chest*, **126**, 1733-1741.

- Gold, B., Cankovic, M., Furtado, L.V., Meier, F. & Gocke, C.D. (2015) Do Circulating Tumor Cells, Exosomes, and Circulating Tumor Nucleic Acids Have Clinical Utility?: A Report of the Association for Molecular Pathology. *J Mol Diagn*.
- Govindan, R., Ding, L., Griffith, M., Subramanian, J., Dees, N.D., Kanchi, K.L., Maher, C.A., Fulton, R., Fulton, L., Wallis, J., Chen, K., Walker, J., McDonald, S., Bose, R., Ornitz, D., Xiong, D., You, M., Dooling, D.J., Watson, M., Mardis, E.R. & Wilson, R.K. (2012) Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell*, **150**, 1121-1134.
- Graus-Porta, D., Beerli, R.R., Daly, J.M. & Hynes, N.E. (1997) ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signaling. *EMBO J*, **16**, 1647-1655.
- Greig, S.L. (2016) Osimertinib: First Global Approval. *Drugs*, **76**, 263-273.
- Hanahan, D. & Weinberg, R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, **144**, 646-674.
- Harris, D.A., Patel, S.H., Gucek, M., Hendrix, A., Westbroek, W. & Taraska, J.W. (2015) Exosomes released from breast cancer carcinomas stimulate cell movement. *PLoS One*, **10**, e0117495.
- Hata, A. & Lieberman, J. (2015) Dysregulation of microRNA biogenesis and gene silencing in cancer. *Sci Signal*, **8**, re3.
- Hofman, V., Ilie, M.I., Long, E., Selva, E., Bonnetaud, C., Molina, T., Vénissac, N., Mouroux, J., Vielh, P. & Hofman, P. (2011) Detection of circulating tumor cells as a prognostic factor in patients undergoing radical surgery for non-small-cell lung carcinoma: comparison of the efficacy of the CellSearch Assay™ and the isolation by size of epithelial tumor cell method. *Int J Cancer*, **129**, 1651-1660.
- Huber, R.M. & Stratakis, D.F. (2004) Molecular oncology--perspectives in lung cancer. *Lung Cancer*, **45 Suppl 2**, S209-213.
- Hynes, N.E. & Lane, H.A. (2005) ERBB receptors and cancer: the complexity of targeted inhibitors. *Nat Rev Cancer*, **5**, 341-354.
- Imielinski, M., Berger, A.H., Hammerman, P.S., Hernandez, B., Pugh, T.J., Hodis, E., Cho, J., Suh, J., Capelletti, M., Sivachenko, A., Sougnez, C., Auclair, D., Lawrence, M.S., Stojanov, P., Cibulskis, K., Choi, K., de Waal, L., Sharifnia, T., Brooks, A., Greulich, H., Banerji, S., Zander, T., Seidel, D., Leenders, F., Ansén, S., Ludwig, C., Engel-Riedel, W., Stoelben, E., Wolf, J., Goparju, C., Thompson, K., Winckler, W., Kwiatkowski, D., Johnson, B.E., Jänne, P.A., Miller, V.A., Pao, W., Travis, W.D., Pass, H.I., Gabriel, S.B., Lander, E.S., Thomas, R.K., Garraway, L.A., Getz, G. & Meyerson, M. (2012) Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell*, **150**, 1107-1120.
- Jackson, R.J. & Standart, N. (2007) How do microRNAs regulate gene expression? *Sci STKE*, **2007**, re1.
- Jahr, S., Hentze, H., Englisch, S., Hardt, D., Fackelmayer, F.O., Hesch, R.D. & Knippers, R. (2001) DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res*, **61**, 1659-1665.
- Jiang, N. & Pisetsky, D.S. (2005) The effect of inflammation on the generation of plasma DNA from dead and dying cells in the peritoneum. *J Leukoc Biol*, **77**, 296-302.
- Kadri, S., Zhen, C.J., Wurst, M.N., Long, B.C., Jiang, Z.F., Wang, Y.L., Furtado, L.V. & Segal, J.P. (2015) Amplicon Indel Hunter Is a Novel Bioinformatics Tool to Detect Large Somatic Insertion/Deletion Mutations in Amplicon-Based Next-Generation Sequencing Data. *J Mol Diagn*, **17**, 635-643.
- Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., Leiserson, M.D., Miller, C.A., Welch, J.S., Walter, M.J., Wendl, M.C., Ley, T.J., Wilson, R.K., Raphael, B.J. & Ding, L. (2013)

- Mutational landscape and significance across 12 major cancer types. *Nature*, **502**, 333-339.
- Kapliniski, L., Andreson, R., Puurand, T. & Remm, M. (2005) MultiPLX: automatic grouping and evaluation of PCR primers. *Bioinformatics*, **21**, 1701-1702.
- Kim, S.T., Sung, J.S., Jo, U.H., Park, K.H., Shin, S.W. & Kim, Y.H. (2013) Can mutations of EGFR and KRAS in serum be predictive and prognostic markers in patients with advanced non-small cell lung cancer (NSCLC)? *Med Oncol*, **30**, 328.
- Kok-Sin, T., Mokhtar, N.M., Ali Hassan, N.Z., Sagap, I., Mohamed Rose, I., Harun, R. & Jamal, R. (2015) Identification of diagnostic markers in colorectal cancer via integrative epigenomics and genomics data. *Oncol Rep*, **34**, 22-32.
- Krebs, M.G., Sloane, R., Priest, L., Lancashire, L., Hou, J.M., Greystoke, A., Ward, T.H., Ferraldeschi, R., Hughes, A., Clack, G., Ranson, M., Dive, C. & Blackhall, F.H. (2011) Evaluation and prognostic significance of circulating tumor cells in patients with non-small-cell lung cancer. *J Clin Oncol*, **29**, 1556-1563.
- Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v1 [q-bio.GN].
- Lo, Y.M., Rainer, T.H., Chan, L.Y., Hjelm, N.M. & Cocks, R.A. (2000) Plasma DNA as a prognostic marker in trauma patients. *Clin Chem*, **46**, 319-323.
- Luo, J., Shen, L. & Zheng, D. (2014) Diagnostic value of circulating free DNA for the detection of EGFR mutation status in NSCLC: a systematic review and meta-analysis. *Sci Rep*, **4**, 6269.
- Ma, M., Zhu, H., Zhang, C., Sun, X., Gao, X. & Chen, G. (2015) "Liquid biopsy"-ctDNA detection with great potential and challenges. *Ann Transl Med*, **3**, 235.
- MacConaill, L.E., Van Hummelen, P., Meyerson, M. & Hahn, W.C. (2011) Clinical implementation of comprehensive strategies to characterize cancer genomes: opportunities and challenges. *Cancer Discov*, **1**, 297-311.
- Mauger, F., Dulary, C., Daviaud, C., Deleuze, J.F. & Tost, J. (2015) Comprehensive evaluation of methods to isolate, quantify, and characterize circulating cell-free DNA from small volumes of plasma. *Anal Bioanal Chem*, **407**, 6873-6878.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. & DePristo, M.A. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*, **20**, 1297-1303.
- McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P. & Cunningham, F. (2010) Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*, **26**, 2069-2070.
- Miller, A.B., Altenburg, H.P., Bueno-de-Mesquita, B., Boshuizen, H.C., Agudo, A., Berrino, F., Gram, I.T., Janson, L., Linseisen, J., Overvad, K., Rasmuson, T., Vineis, P., Lukanova, A., Allen, N., Amiano, P., Barricarte, A., Berglund, G., Boeing, H., Clavel-Chapelon, F., Day, N.E., Hallmans, G., Lund, E., Martinez, C., Navarro, C., Palli, D., Panico, S., Peeters, P.H., Quirós, J.R., Tjønneland, A., Tumino, R., Trichopoulou, A., Trichopoulos, D., Slimani, N. & Riboli, E. (2004) Fruits and vegetables and lung cancer: Findings from the European Prospective Investigation into Cancer and Nutrition. *Int J Cancer*, **108**, 269-276.
- Paci, M., Maramotti, S., Bellesia, E., Formisano, D., Albertazzi, L., Ricchetti, T., Ferrari, G., Annessi, V., Lasagni, D., Carbonelli, C., De Franco, S., Brini, M., Sgarbi, G. & Lodi, R. (2009) Circulating plasma DNA as diagnostic biomarker in non-small cell lung cancer. *Lung Cancer*, **64**, 92-97.
- Padrik, P. & Everaus, H. (2013) *Onkoloogia õpik*, Tartu Ülikooli Kirjastus, Tartu.
- Page, K., Guttery, D.S., Zahra, N., Primrose, L., Elshaw, S.R., Pringle, J.H., Blighe, K., Marchese, S.D., Hills, A., Woodley, L., Stebbing, J., Coombes, R.C. & Shaw, J.A.

- (2013) Influence of plasma processing on recovery and analysis of circulating nucleic acids. *PLoS One*, **8**, e77963.
- Palli, D., Saieva, C., Grechi, D., Masala, G., Zanna, I., Barbaro, A., Decarli, A., Munnia, A. & Peluso, M. (2004) DNA bulky adducts in a Mediterranean population correlate with environmental ozone concentration, an indicator of photochemical smog. *Int J Cancer*, **109**, 17-23.
- Pope, C.A., Burnett, R.T., Thun, M.J., Calle, E.E., Krewski, D., Ito, K. & Thurston, G.D. (2002) Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA*, **287**, 1132-1141.
- Rekhtman, N., Ang, D.C., Sima, C.S., Travis, W.D. & Moreira, A.L. (2011) Immunohistochemical algorithm for differentiation of lung adenocarcinoma and squamous cell carcinoma based on large series of whole-tissue sections with validation in small specimens. *Mod Pathol*, **24**, 1348-1359.
- Riese, D.J. & Stern, D.F. (1998) Specificity within the EGF family/ErbB receptor family signaling network. *Bioessays*, **20**, 41-48.
- Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. & Mesirov, J.P. (2011) Integrative genomics viewer. *Nat Biotechnol*, **29**, 24-26.
- Rolfo, C., Castiglia, M., Hong, D., Alessandro, R., Mertens, I., Baggerman, G., Zwaenepoel, K., Gil-Bazo, I., Passiglia, F., Carreca, A.P., Taverna, S., Vento, R., Santini, D., Peeters, M., Russo, A. & Pauwels, P. (2014) Liquid biopsies in lung cancer: the new ambrosia of researchers. *Biochim Biophys Acta*, **1846**, 539-546.
- Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., Yan, H., Gazdar, A., Powell, S.M., Riggins, G.J., Willson, J.K., Markowitz, S., Kinzler, K.W., Vogelstein, B. & Velculescu, V.E. (2004) High frequency of mutations of the PIK3CA gene in human cancers. *Science*, **304**, 554.
- Schwarzenbach, H., Hoon, D.S. & Pantel, K. (2011) Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer*, **11**, 426-437.
- Sequist, L.V., Waltman, B.A., Dias-Santagata, D., Digumarthy, S., Turke, A.B., Fidias, P., Bergethon, K., Shaw, A.T., Gettinger, S., Cosper, A.K., Akhavanfard, S., Heist, R.S., Temel, J., Christensen, J.G., Wain, J.C., Lynch, T.J., Vernovsky, K., Mark, E.J., Lanuti, M., Iafrate, A.J., Mino-Kenudson, M. & Engelman, J.A. (2011) Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Sci Transl Med*, **3**, 75ra26.
- Shigematsu, H., Lin, L., Takahashi, T., Nomura, M., Suzuki, M., Wistuba, I.I., Fong, K.M., Lee, H., Toyooka, S., Shimizu, N., Fujisawa, T., Feng, Z., Roth, J.A., Herz, J., Minna, J.D. & Gazdar, A.F. (2005) Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J Natl Cancer Inst*, **97**, 339-346.
- Siegel, R.L., Miller, K.D. & Jemal, A. (2016) Cancer statistics, 2016. *CA Cancer J Clin*, **66**, 7-30.
- Sozzi, G., Conte, D., Leon, M., Ciricione, R., Roz, L., Ratcliffe, C., Roz, E., Cirenei, N., Bellomi, M., Pelosi, G., Pierotti, M.A. & Pastorino, U. (2003) Quantification of free circulating DNA as a diagnostic marker in lung cancer. *J Clin Oncol*, **21**, 3902-3908.
- Stroun, M., Lyautey, J., Lederrey, C., Olson-Sand, A. & Anker, P. (2001) About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. *Clin Chim Acta*, **313**, 139-142.
- Su, Z., Dias-Santagata, D., Duke, M., Hutchinson, K., Lin, Y.L., Borger, D.R., Chung, C.H., Massion, P.P., Vnencak-Jones, C.L., Iafrate, A.J. & Pao, W. (2011) A platform for rapid detection of multiple oncogenic mutations with relevance to targeted therapy in non-small-cell lung cancer. *J Mol Diagn*, **13**, 74-84.

- Subramanian, J., Velcheti, V., Gao, F. & Govindan, R. (2007) Presentation and stage-specific outcomes of lifelong never-smokers with non-small cell lung cancer (NSCLC). *J Thorac Oncol*, **2**, 827-830.
- Swaminathan, R. & Butt, A.N. (2006) Circulating nucleic acids in plasma and serum: recent developments. *Ann N Y Acad Sci*, **1075**, 1-9.
- Szpechcinski, A., Rudzinski, P., Kupis, W., Langfort, R., Orlowski, T. & Chorostowska-Wynimko, J. (2016) Plasma cell-free DNA levels and integrity in patients with chest radiological findings: NSCLC versus benign lung nodules. *Cancer Lett*, **374**, 202-207.
- Tong, L., Spitz, M.R., Fueger, J.J. & Amos, C.A. (1996) Lung carcinoma in former smokers. *Cancer*, **78**, 1004-1010.
- Travis, W.D., Brambilla, E., Nicholson, A.G., Yatabe, Y., Austin, J.H., Beasley, M.B., Chirieac, L.R., Dacic, S., Duhig, E., Flieder, D.B., Geisinger, K., Hirsch, F.R., Ishikawa, Y., Kerr, K.M., Noguchi, M., Pelosi, G., Powell, C.A., Tsao, M.S., Wistuba, I. & Panel, W. (2015) The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. *J Thorac Oncol*, **10**, 1243-1260.
- Vogelstein, B. & Kinzler, K.W. (2004) Cancer genes and the pathways they control. *Nat Med*, **10**, 789-799.
- Wang, A., Kubo, J., Luo, J., Desai, M., Hedlin, H., Henderson, M., Chlebowski, R., Tindle, H., Chen, C., Gomez, S., Manson, J.E., Schwartz, A.G., Wactawski-Wende, J., Cote, M., Patel, M.I., Stefanick, M.L. & Wakelee, H.A. (2015) Active and passive smoking in relation to lung cancer incidence in the Women's Health Initiative Observational Study prospective cohort. *Ann Oncol*, **26**, 221-230.
- Wang, S., An, T., Wang, J., Zhao, J., Wang, Z., Zhuo, M., Bai, H., Yang, L., Zhang, Y., Wang, X., Duan, J., Wang, Y., Guo, Q. & Wu, M. (2010) Potential clinical significance of a plasma-based KRAS mutation analysis in patients with advanced non-small cell lung cancer. *Clin Cancer Res*, **16**, 1324-1330.
- Wong, D., Moturi, S., Angkachatchai, V., Mueller, R., DeSantis, G., van den Boom, D. & Ehrich, M. (2013) Optimizing blood collection, transport and storage conditions for cell free DNA increases access to prenatal testing. *Clin Biochem*, **46**, 1099-1104.
- Xie, W., Qin, W., Kang, Y., Zhou, Z. & Qin, A. (2016) MicroRNA-340 Inhibits Tumor Cell Proliferation and Induces Apoptosis in Endometrial Carcinoma Cell Line RL 95-2. *Med Sci Monit*, **22**, 1540-1546.
- Yarden, Y. & Sliwkowski, M.X. (2001) Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol*, **2**, 127-137.
- Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S., Gil, L., Girón, C.G., Gordon, L., Hourlier, T., Hunt, S.E., Janacek, S.H., Johnson, N., Juettemann, T., Keenan, S., Lavidas, I., Martin, F.J., Maurel, T., McLaren, W., Murphy, D.N., Nag, R., Nuhn, M., Parker, A., Patricio, M., Pignatelli, M., Rahtz, M., Riat, H.S., Sheppard, D., Taylor, K., Thormann, A., Vullo, A., Wilder, S.P., Zadissa, A., Birney, E., Harrow, J., Muffato, M., Perry, E., Ruffier, M., Spudich, G., Trevanion, S.J., Cunningham, F., Aken, B.L., Zerbino, D.R. & Flicek, P. (2016) Ensembl 2016. *Nucleic Acids Res*, **44**, D710-716.
- Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S. & Madden, T.L. (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, **13**, 134.
- Yu, H.A., Arcila, M.E., Rekhtman, N., Sima, C.S., Zakowski, M.F., Pao, W., Kris, M.G., Miller, V.A., Ladanyi, M. & Riely, G.J. (2013) Analysis of tumor specimens at the time of acquired resistance to EGFR-TKI therapy in 155 patients with EGFR-mutant lung cancers. *Clin Cancer Res*, **19**, 2240-2247.
- Zhang, Z., Xiao, Y., Zhao, J., Chen, M., Xu, Y., Zhong, W., Xing, J. & Wang, M. (2016) Relationship between circulating tumour cell count and prognosis following

chemotherapy in patients with advanced non-small-cell lung cancer. *Respirology*, **21**, 519-525.

Kasutatud veebiaadressid

cancer.sanger.ac.uk

<http://pxweb.tai.ee/esf/pxweb2008/dialog/statfile2.asp>

<http://www.cancer.org/cancer/lungcancer-non-smallcell/detailedguide/non-small-cell-lung-cancer-what-is-non-small-cell-lung-cancer>

<http://www.idtdna.com/Scitools>

http://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_trusight_tumor.pdf

http://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf

<http://www.ncbi.nlm.nih.gov/gene>

<http://www.tai.ee/et/tegevused/registrid/vahiregister>

LISAD

Lisa 1. Modifitseeritud NucleoSpin[®] Plasma XS high sensitivity protokoll

- *1. Prepare sample: add 720 µL plasma to a microcentrifuge tube (2 ml, not provided).
- *1a Proteinase K treatment: add 60 µL Proteinase K to the plasma sample, mix, and incubate at 37 °C for 10 min.
- *2. Adjust DNA binding conditions: add 1080 µL Buffer BB.
- 3. Mix sample: invert the tube 3x and vortex for 3 s. Centrifuge the tube briefly to clean the lid.
- 4. Bind DNA: for each sample, load the mixture (600 µL) to a NucleoSpin[®] Plasma XS Column placed in a Collection Tube (2 mL). Centrifuge at 4500 rpm for 30 s, increase centrifuge force to 12,000 rpm for further 60 s. Discard flow-through. Discard Collection Tube with flow-through and place column into new Collection Tube (provided).
- 5. Wash and dry silica membrane:
 - 1st wash: pipette 500 µL Buffer WB onto the NucleoSpin[®] Plasma XS Column. Centrifuge for 30 s at 12,000 rpm. Discard flow-through.
 - 2nd wash: add 250 µL Buffer WB to the NucleoSpin[®] Plasma XS Column. Centrifuge for 3 min at 12,000 rpm. Discard Collection Tube with flow-through and place the column into a 1.5 mL microcentrifuge tube for elution (not provided).
- 6. Elute DNA: add 20 µL Elution Buffer to the NucleoSpin[®] Plasma XS Column. Centrifuge for 30 s at 12,000 rpm.
- *7. Removal of residual ethanol: Incubate elution fraction with open lid for 17 min at 75 °C.

* tähistatud punkte on modifitseeritud.

Lisa 2. Modifitseeritud QIAamp[®] CNA protokoll 6 ml plasmas rakuvaba DNA eraldamiseks

- *1. Pipet 400 µl QIAGEN Proteinase K into a 50 ml centrifuge tube (not provided).
- *2. Add 6 ml plasma to the tube.
- *3. Add 3,25 ml Buffer ACL. Close the cap and mix by pulse-vortexing for 30 s.

4. Incubate at 60 °C for 30 min.
5. Place the tube back on the lab bench and unscrew the cap.
- *6. Add 7 ml Buffer ACB to the lysate in the tube. Close the cap and mix thoroughly by pulse-vortexing for 15–30 s.
7. Incubate the lysate–Buffer ACB mixture in the tube for 5 min on ice.
8. Insert the QIAamp Mini column into the VacConnector on the QIAvac 24 Plus. Insert a 20 ml tube extender into the open QIAamp Mini column.
9. Carefully apply the lysate–Buffer ACB mixture from step 7 into the tube extender of the QIAamp Mini column. Switch on the vacuum pump (800 – 900 mbar). When all lysates have been drawn through the columns completely, switch off the vacuum pump and release the pressure to 0 mbar. Carefully remove and discard the tube extender.
10. Apply 600 µl Buffer ACW1 to the QIAamp Mini column. Leave the lid of the column open, and switch on the vacuum pump (800 – 900 mbar). After all of Buffer ACW1 has been drawn through the QIAamp Mini column, switch off the vacuum pump and release the pressure to 0 mbar.
- *11. Apply 700 µl Buffer ACW2 to the QIAamp Mini column. Leave the lid of the column open, and switch on the vacuum pump (800 – 900 mbar). After all of Buffer ACW2 has been drawn through the QIAamp Mini column, switch off the vacuum pump and release the pressure to 0 mbar.
- *12. Apply 700 µl of ethanol (96–100%, not provided) to the QIAamp Mini column. Leave the lid of the column open, and switch on the vacuum pump (800 – 900 mbar). After all of ethanol has been drawn through the spin column, switch off the vacuum pump and release the pressure to 0 mbar.
13. Close the lid of the QIAamp Mini column. Remove it from the vacuum manifold, and discard the VacConnector. Place the QIAamp Mini column in a clean 2 ml collection tube, and centrifuge at full speed (13,000 x g) for 3 min.
14. Place the QIAamp Mini Column into a new 2 ml collection tube. Open the lid, and incubate the assembly at 56 °C for 10 min to dry the membrane completely.

*15. Place the QIAamp Mini column in a clean 1,5 ml elution tube (provided) and discard the 2 ml collection tube from step 14. Carefully apply 60 µl of (warmed) Buffer AVE to the center of the QIAamp Mini membrane. Close the lid and incubate at room temperature for 3 min.

*16. Centrifuge in a microcentrifuge at full speed (13,000 x g) for 1 min to elute the nucleic acids. Repeat the elution step.

* tähistatud punkte on modifitseeritud.

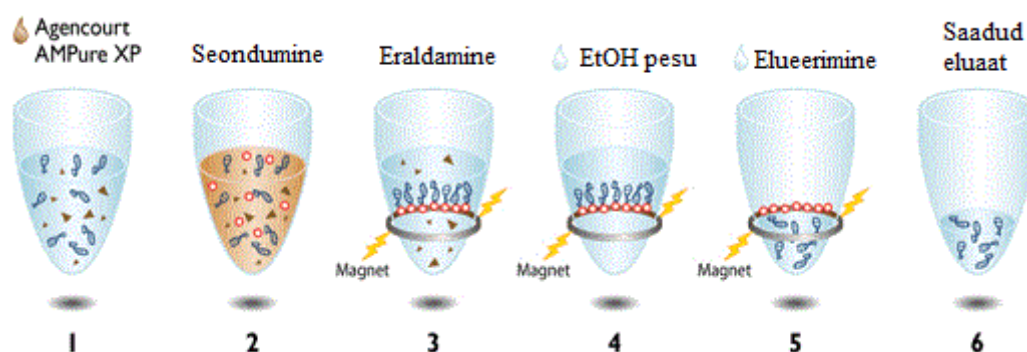
Lisa 3. PCR produktide puhastamise protokoll Agencourt AMPure XP magnetkerakestega (Tartu Ülikooli Eesti Geenivaramu tuumiklabor, Eesti)

1. Too Agencourt AMPure XP (Beckman Coulter, Inc., USA) magnetkerakesed toatemperatuurile ning seejärel sega korralikult.
2. Lisa PCR-i reaktsioonisegule 20 µl* AMPure XP magnetkerakesi ja raputa 1800 rpm (*revolutions per minute*) 2 minutit.
3. DNA seondumiseks kerakestele inkubeeri segu 5 minutit toatemperatuuril.
4. Aseta PCR-i plaat magnethoidjale. Kui supernatant on muutunud selgeks, eemalda ettevaatlikult vedelik.
5. Jäta plaat magnethoidjale ning lisa 200 µl värskelt valmistatud 80% EtOH lahust. Inkubeeri 30 sekundit ning seejärel eemalda kogu vedelik. Korda etanooliga pesu samal viisil. Eemalda peenikest pipetiotsa kasutades kogu etanool.
6. Lase kerakestel magnethoidjal ilma kaaneta 10 minutit kuivada.
7. Eemalda kuivanud kerakestega plaat magnethoidjalt ja lisa 52,5 µl** 10 mM Tris pH 8,5 puhvrit. Raputa plaati 1800 rpm 2 minutit.
8. DNA elueerimiseks kerakestelt inkubeeri lahust toatemperatuuril 2 minutit.
9. Aseta PCR-i plaat magnethoidjale. Kui vedelik on muutunud selgeks, tõsta ettevaatlikult 50 µl*** eluaati uuele plaadile.

* Amplikon PCR-i produktidele tuleb lisada AMPure XP magnetkerakesi 20 µl, kuid indeks PCR-i produktidele 56 µl.

** Amplikon PCR-i produktidele tuleb lisada Tris puhvrit 52,5 µl, kuid indeks PCR-i produktidele 27,5 µl.

*** Amplikon PCR-i produktide puhul tuleb tõsta 50 µl eluaati uuele plaadile, kuid indeks PCR-i produktide puhul 25 µl.



Lisa 3.1. Agencourt AMPure XP magnetkerakestega puhastamise põhimõte (Beckman Coulter, Inc., USA).

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Madli Tamm (sünnikuupäev: 06.05.1990),

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose
„Kasvajaseoseliste mutatsioonide määramine kopsuvähiga uuritavate rakuvabast DNA-st“,
mille juhendajad on MD, PhD Neeme Tõnisson, PhD Paula Ann Kivistik ja MSc Mart Kals
 - 1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
 - 1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 27.05.2016